OIPE
AUG 0 1 2002
PATENT & TRADEMARK OFFICE
3C95

SGM 6934.1
PATENT

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Application of Ward et al.
Serial No.: 09/610,935
Filed: July 6, 2000
For: TARGET REAGENTS THAT ENHANCE REACTION-PRODUCT ANALYSIS
Examiner: Bradley L. Sisson
Confirmation No. 5148

Art Unit: 1655

July 29, 2002

Declaration of Brian W. Ward Under 37 C.F.R. 1.132

TO THE COMMISSIONER OF PATENTS AND TRADEMARKS
Washington, D.C. 20231

I, Brian W. Ward, declare as follows:

1. I am a citizen of the United States of America, and I reside in St. Louis County, Missouri.

2. I am currently employed at Sigma-Aldrich Co. as a research scientist. I have a substantial amount of formal education, training and experience in biochemical and biological research, and particularly, in the field of molecular biology, as evidenced by my curriculum vitae, attached hereto as Exhibit B.

3. I am a co-inventor of the above-referenced patent application, and therefore, am familiar with the contents and substance of the specification and presently-pending claims.

4. The present invention is directed toward novel compositions comprising a polymerase for an *ex-vivo* enzymatic reaction in which a nucleic acid polymer product complementary to a nucleic acid polymer template is prepared, and a tracer dye compatible with the polymerase, the composition being substantially free of the nucleic acid polymer template. Particular embodiments of the invention comprise various composition densities, various optical densities, various dyes, including, for example, acid violet 5, acid red 1, and combinations thereof, and various polymerases, including, for example, Taq polymerase.

5. The compositions of the present invention are used for an enzymatic reaction and enhance reaction-product analysis. Specifically, certain preferred embodiments of the invention provide compositions of comprising components which facilitate subsequent chromatographic or electrophoretic analysis.

6. In order to exemplify the invention, one particular embodiment of the invention is demonstrated in Example 1, specification p. 18, line 6 et seq. This particular embodiment demonstrates the use of Taq polymerase as the polymerase and a combination of 80% acid red 1 and 20% acid violet 5 as the tracer dye.

7. Taq polymerase is a nucleotidyl transferase – *i.e.*, an enzyme of the transferase group that catalyzes the transfer of a nucleotidyl group, such as for example, DNA polymerase or RNA polymerase. Such enzymes are classified based upon their template (the strand of DNA or RNA that specifies the nucleotide sequence of the newly synthesized product strand) and reaction products (the resulting DNA or RNA strand that is newly synthesized). Of particular importance with respect to the present application are the following two classes of polymerases:

   a. DNA-dependent DNA polymerases: employed in all cells for replication and repair of the DNA.

   b. RNA-dependent DNA polymerases: employed by retro viruses in order to copy RNA into DNA (these polymerases commonly called reverse transcriptases (RT)).

8. As of the filing date of the present application, July 6, 2000, it was common knowledge, as demonstrated by this Declaration and the references cited herein, that DNA polymerases share commonly conserved structural and functional characteristics. Therefore, as one of ordinary skill in the art, the disclosure and use of Taq polymerase in Example 1 of the present application would be sufficiently descriptive to confer to me and others that the Applicants of the present invention were in possession not only of the particular embodiment disclosed above, but of all such embodiments comprising a DNA polymerase, and more particularly a thermostable polymerase, and a tracer dye, and more particularly an anionic dye. Specifically, because of the high knowledge and skill in the art regarding the structure, function, and use of polymerases, and in particular thermostable polymerases, one would be well aware of the fact that the Applicants were in possession of all of the embodiments of the invention as claimed at the time of filing the present application.

9. Most DNA polymerases possess various structural and functional characteristics in common. These characteristics are highly conserved, even among completely unrelated polymerases. The presence of these common characteristics creates a rather unvarying overall structure among DNA polymerases. Furthermore, they impart upon the polymerases a substantially uniform mechanism of action.

10. Specifically, the first crystal structure of a DNA polymerase (the Klenow fragment of *E. coli* Pol I) provided priceless insights into the structure of polymerases, and particularly,

how polymerases are designed to catalyze DNA replication. *See*, Hingorani, M.M. and O'Donnell, M., "DNA Polymerase Structure and Mechanisms of Action", found at http://bentham.org/coc-sample/o-donnell/o-donnell.htm, and attached hereto as Exhibit C. As described in the reference, the structure of the first Klenow fragment is often said to resemble a half-open right hand. That is, the structure possess a "palm" subdomain (a subdomain forming a cleft), flanked by a "fingers" subdomain and a "thumb" subdomain. This structure is demonstrated in Figure 1 of the Hingorani and O'Donnell reference and is also attached hereto as Exhibit D. This particular shape is by far the most prominent feature common to all known polymerase structures. Specifically, several other polymerase structures determined subsequent to the Klenow fragment have demonstrated that, despite some differences in amino acid sequences and structures among enzymes from different families, several features *essential* for polymerase activity, including the particular shape outlined above, are highly conserved. Examples of this include, but are not limited to, members of the Pol I family of polymerases (Klenow fragment, *Thermus aquaticus* pol, *T7* DNA pol, and *Bacillus stearothermophilus* pol); members of the Pol II family of polymerases (RB69 pol, *D. tok* Pol); members of the family X of polymerases (pol β); and members of the RT family (HIV-RT). Such commonalities are also present within the family of thermostable DNA polymerases, including for example Taq (*Thermus aquaticus*), Klentaq (*Thermus aquaticus*, N-terminal deletion mutant), Vent/Tli (*Thermococcus litoralis*), Deep Vent (*Pyrococcus* species *GB-D*), Pfu (*Purococcus furiosus*), Replinase (*Thermus flavis*), KOD (*Pyrococcus kodakaraensis* KOD1), Tth (*Thermus thermophilus* HB-8), and Tfl (*Thermus flavus*).

11. The three subdomains (palm, fingers, and thumb) act synergistically to hold the primer template DNA, as well as to position the incoming dNTP for incorporation into the elongating DNA strand. The palm subdomain contains the catalytic site wherein the chemistry of nucleotidy transfer occurs. The fingers subdomain interacts with the template DNA strand and the incoming dNTP, properly positioning the two. The thumb subdomain is primarily responsible for binding the duplex DNA in a sequence-independent manner along the minor groove of the DNA. *See*, Hingorani and O'Donnell.

12. Attached hereto as Exhibit E is a list of websites discussing polymerase structure. Information contained therein corroborates the statements contained in this declaration, as well as the statements in the Hingorani and McDonnell reference and the additional references cited below in paragraph 18. Furthermore, the statements contained in the Hingorani and McDonnell reference cited above have recently been independently reviewed in Brautigam, C.A. and Steitz, T.A. (1998) *Curr Opin Struct Biol*, 8:54-63 and Sousa, R., (1996) *Trends Biochem Sci*, 21:186-90.

13. As with structure, the reaction mechanism displayed by polymerases, including Taq polymerase and other thermostable polymerases, is universal among the polymerases. Specifically, all polymerases follow the mechanism for T7 DNA polymerase, a figure

illustrating said mechanism demonstrated in Figure 2 of Hingorani and O'Donnell and attached hereto as Exhibit F.

14. The structural basis of this catalysis (*i.e.*, reaction mechanism) relies upon a two metal center mechanism. Figure 3 of Hingorani and O'Donnell, attached hereto as Exhibit G, illustrates the catalytic site, including illustrations of the T7 DNA polymerase active site (A. of Exhibit G) and the two-metal ion mechanism for nucleotidyl transfer (B. of Exhibit G).

15. As specifically stated in Hingorani and McDonnell (and as supported by Alba, Jager and Prata, Brautigam and Steitz, and Sousa (citations listed in paragraph 18 below)), the catalysis reaction occurs via a nucleotidyl transfer reaction wherein the 3'-hydroxyl group of the primer attacks the alpha-phosphate of the incoming dNTP. This occurs in the catalytic site located within the palm domain of the polymerase. The active site, as illustrated in Exhibit G, contains several acidic and polar amino acid residues in addition to two metal cations. These two cations, usually consisting of Mg^{2+} ions, are essential for the catalysis mechanism. Additionally, two aspartate residues are absolutely conserved between the polymerase families. These residues provide the carboxylate oxygen molecules that ligate the two metal ions. Ion A (as designated in A. of Exhibit G) is found near the 3'-hydroxyl group of the DNA primer and the alpha-phosphate of the incoming dNTP. In this position, ion A is situated to lower the pKa of the hydroxyl group and to facilitate formation of the alkoxide anion, which can thereby initiate nucleophilic attack on the alpha-phosphate of the dNTP. Ion A also ligates the alpha-phosphate of dNTP, thereby aiding in the stabilization of the pentacovalent transition-state formed during the reaction. Specifically, ion A acts to stabilize the transition state (the state wherein an alpha phosphorous is penta-coordinated to five oxygen molecules, resulting in the transition of the alpha phosphate from a tetrahedral structure to a planar-bipyramidal structure). *See also*, Slickers, Peter, "DNA Polymerases," http://www.imb-jena.de/ ~slickers/polymerases/dna-polymerases.html

16. As further stated in Hingorani and McDonnell, metal ion B (as designated in B. of Exhibit G) ligates oxygen groups in all three of the phosphate groups of the dNTP. It is believed that this aids in the alignment of the triphosphate moiety for attack by the 3'-hydroxyl, in addition to stabilizing the charge on the transition state and aiding in the positioning of incoming mononucleotides. Additional polar residues located within the active site, and possibly ion B as well, aid in the stabilization of the charged pyrophosphate group as it dissociates from the polymerase upon completion of the nucleotidyl transfer.

17. The catalytic mechanism described above, as well as the geometry of the active site, are highly conserved among the polymerases. This conservation occurs even among the

completely unrelated polymerases such as the T7 DNA polymerase and the Pol β polymerase.

18. In addition to the references cited herein, the above statements regarding DNA polymerase structure and reaction mechanism are supported by the following references, attached hereto as Exhibit C:

a. Alba, M. "Replicative DNA polymerases." *Genome Biol* (2001) 2: REVIEWS3002.1.

b. Brautigam C.A. and Steitz, T.A. "Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes." *Curr Opin Struct Biol* (1998) 8:54-63.

c. Hingorani1, M.M. and O'Donnell, M. "DNA Polymerase Structure and Mechanisms of Action." http://www.bentham.org/coc-sample/o-donnell/o-donnell.htm

d. Jager J. and Pata, J.D. "Getting a grip: polymerases and their substrate complexes" *Curr Opin Struct Biol* (1999) 9:21-8.

e. Slickers, Peter. "DNA Polymerases," http://www.imb-jena.de/~slickers/polymerases/dna-polymerases.html

f. Sousa R. "Structural and mechanistic relationships between nucleic acid polymerases." *Trends Biochem Sci* (1996) 21, 186-90.

19. The level of knowledge and skill in the field with respect to DNA polymerases, as can be seen from the above cited references, is quite high. That is to say, those of skill in the art are aware of the great similarities among various polymerases. Furthermore, those of even *moderate* skill in the art have the knowledge and the skill sufficient to enable them to use polymerases, as such knowledge and skill are taught to students in undergraduate molecular biology classes and are practiced by undergraduate students in their lab classes. Furthermore, this knowledge and skill, combined with the disclosure of Example 1, would be sufficient to convey to those of even moderate skill in the art that Applicants were in possession of not only a composition comprising Taq polymerase as the polymerase component of the composition, but of a composition comprising any DNA polymerase.

20. Additionally, the dyes used in the particular embodiments of the present invention are representative of numerous dyes that can be used with similar success. Particularly, numerous other anionic dyes would be suitable for use in the present invention as outlined in the specification, and particularly, in Example 1.

21. As outlined in Example 1, anionic dyes were selected for screening because they would work best with the invention as disclosed. Specifically, anionic dyes have the least likelihood of interacting with the nucleic acids during the polymerization reaction. This is as a result of the charge-charge repulsion. Specifically, since the nucleic acid products

of PCR are highly anionic, they are therefore applied to the agarose gel near the cathode and move toward the anode as the electrophoresis progresses. Therefore, to be of use as a tracer in electrophoresis, the dye is also preferably anionic as well. In contrast, organic cations will almost certainly interact with polyanions such as polynucleic acids. Neutral dyes are also of little use due to their limited aqueous solubility.

22. Furthermore, about one-half of the anionic dyes were thereafter eliminated from consideration by the mere fact that they were not of a red color. Specifically, Applicants demonstrated in Example 1 only dyes that were red. This color was selected for no reason other than aesthetics. The assignee of the present application, Sigma-Aldrich Co., typically selects products displaying red or a particular shade of red, as the color red is often associated with the Sigma-Aldrich Co. and many of the marks it has registered.

23. The use of a red dye conferred no particular advantage on the embodiment of Example 1, and there is no reason why dyes of other colors would not be described by and encompassed within the present invention.

24. The dyes were further eliminated from consideration on the basis of characteristic such as ethanol precipitation, solid phase extraction, PCR toxicity, and ligation/transformation toxicity. These characteristics were chosen based upon the particularities of Example 1. However, based upon the structural and functional characteristics of Taq polymerase (the polymerase used in Example 1) as well as knowledge and skill regarding anionic dyes, it is clear that numerous other anionic dyes would also work well. Included within this group are the dyes initially eliminated solely because they were not red in color. More specifically, other dyes would work well with DNA polymerases. Artisans practicing the invention would be able to determine, based upon the description and the general knowledge within the art, which dyes would be ideal for the particular embodiment being practiced without undue experimentation.

25. In terms of the generalized polymerase mechanism shown above (Exhibit F), dyes of the present invention would not inhibit polymerization by reducing the concentration of DNA through the creation of a DNA-dye complex. Additionally, as demonstrated by the present application, dyes of the present invention do not inhibit polymerization by other mechanisms, such as mechanisms in which the dye binds to the enzyme. Specifically, as demonstrated in the present application, screening a large number of dyes showed that a subset did not adversely bind to Taq DNA polymerase in such a manner as to inhibit polymerization. Therefore, instructive of the present application is the fact that dyes of the present invention do not interact with the polymerase by binding to its structural features.

26. This mode of inhibition was anticipated and/or found to be active or passive in nature. In terms of active inhibition, the dye could competitively bind at one of the catalytic

subdomains thereby inhibiting the binding of key reaction components. It is conceivable that an anionic dye could competitively bind the enzyme at one or more of the template, the primer, or the nucleotide binding surfaces. As demonstrated in the present application, specification p. 28, line 13 through p. 29, line 35, it was found that dyes passively inhibit the enzyme by coordination of the required metal ion. This was, therefore, compensated for by making the dye metal the metal required by the polymerase, followed by optimization of free metal concentration.

27. It was anticipated that the same process would be followed for other polymerase/dye combinations. Specifically, it was believed that the successful use of anionic dyes as described in the present application would not only enable another to use the specific dyes disclosed, but also would enable one to use other anionic dyes that possess similar properties, as the dyes of the specific embodiments are indicative of the characteristics of other dyes that could be used with equally satisfactory results. Likewise, it is asserted that the structural and functional characteristics of the dyes of Example 1, as described therein, would clearly convey to a person of skill in the art that Applicants were in possession of the broader invention as claimed – *i.e.*, wherein the dye tracer component of a composition of the invention comprised an anionic dye that does not interfere with polymerase activity. The description of the dyes used in the particular embodiments of the present invention would indicate to one of skill in the art that Applicants were in possession of the invention as claimed at the time of filing.

28. I further declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further, that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under 18 U.S.C. 1001, and that such willful false statements may jeopardize the validity of the application or any patent issuing thereon.

Brian W. Ward

7/29/02
Date

EXHIBIT B

BRIAN WARD, Ph.D.
1640 Jonquil Dr.
Webster Groves, MO 63119
bwardhome@rocketmail.com
(314) 962-0295 (H)
(314) 771-5765 x3657 (W)

EMPLOYMENT HISTORY

9/88-present. Sigma_Aldrich Inc., St. Louis, MO. Currently technical lead and manager for Gene Expression and Analysis, Sequencing and Labeling, Hybridization and Detection. Developed REDTaq DNA polymerase for which a provisional patent was filed (see Patents section). Was in part responsible for the conception and development of the *Exoclone* technology. This technology obviates the need for restriction digestion of PCR products to yield sticky ended sequences that would be used in a subsequent ligation reaction. Initiated the development of other novel ligation technologies.

Initiated a product line aimed at serving the nucleic acid sequence recognition research community. Most notable accomplishment is the invention of Type IIS Restriction Enzyme Footprinting (see research synopsis A for additional details). Established collaborative/cooperative relationships with internationally renowned scientists resulting in commercialization of the chemical footprinting reagent methidiumpropylEDTA and the triplex specific intercalator BePI. Some of these relationships are also yielding publications.

Was one of two scientists involved in the cloning and expression of commercially significant proteins in E. Coli (see research synopsis B).

Developed and maintained oligonucleotide, DNA sequencing and electrophoresis products. New products included: stop linkers (oligonucleotide cassettes used to introduce stop codons into all reading frames of an expression vector), Maxam and Gilbert sequencing kit and electrophoresis convenience products.

1/87-9/88. Biochemical Development, Sigma Chemical Company, St. Louis, MO.
Developed a novel synthesis for meso-isothiocyanato octaethylporphyrin palladium (II), a molecule that could potentially be exploited as a phosphorescent label. Scaled up the synthesis of 3'-azido-3'deoxythymidine (AZT). Produced/synthesized an assortment of department specific products.

9/84-12/86. Department of Chemistry, Syracuse University, Syracuse, NY.
Developed footprinting protocols and techniques to measure drug binding constants to DNA's from natural sources (see research synopsis C). Collaborated with Bristol-Myers scientists on a number of projects. Wrote a Bristol-Myers funded research grant to synthesize a potential sequence specific DNA cleaving molecule.

9/78-8/84. Department of Chemistry, Michigan State University, East Lansing, MI.
Designed, synthesized and performed physical chemical experiments on heme model systems (see research synopsis D). Taught undergraduate laboratory and recitation courses. Merit Level Teaching Award was received for outstanding performance.

4/83-10/83 Department of Engineering, Michigan State University, East Lansing, MI.
Acted as a consultant to determine the practicality of using photochroic dyes to study fluid flow dynamics. The known, reversibly photochroic dyes owed their photochromogenic property to a photo induced zwitterionic charge transfer complex. Since the colored zwitterion form is a dominant species in polar solvents, non-polar (expensive and likely flammable) solvents would have been required for the

studies. Due to the large volumes of solvent envisioned, it was decided that the proposed experiments were not practical.

ADMINISTRATIVE RESPONSIBILITIES

Manage Gene Expression and Analysis, Sequencing and Labeling, Hybridization and Detection Groups.
Instigated and organized departmental/company seminars.
Interviewed applicants for R&D positions.
Administered grant funds while principle investigator was on sabbatical.

EDUCATION

Post-doctoral, Biophysical Chemistry, Syracuse University, Syracuse, NY (1986)
Ph.D., Organic Chemistry, Michigan State University, East Lansing, MI (1984)
B.S., Chemistry (major), Biology (minor), Mathematics (minor), Western Michigan University, Kalamazoo, Michigan (1977)

HONORS AND AWARDS

Senior Chemist Dow Fellowship Amway Fellowship Merit Level Teaching Award

PROFESSIONAL MEMBERSHIPS

American Association for the Advancement of Science
American Chemical Society

REFERENCES

Available upon request.

Publications
Brian Ward, Ph.D.

1. "Nonbinding Steric Effect on CO and O2 Binding to Hemes. Kinetics of Ligand Binding in Iron-Copper Cofacial Diporphyrins and Strapped Hemes"; B. Ward, C.K. Chang and C.-B. Wang, J. Am. Chem. Soc., 103, 5236 (1981).

2. "Cytochrome Oxidase Models. A Magnetically Coupled Heme-Copper Complex", C.K. Chang, M.S. Koo and B. Ward, J. Chem. Soc. Chem. Commun., 716 (1982).

3. "A Convenient Photochemical Method for Reduction of Ferric Hemes", B. Ward and C.K. Chang, Photochem. Photobiol., 35, 757 (1982).

4. "Red-shifts in the Optical Spectra of Porphyrin Schiff's Bases Upon Protonation", B. Ward, P.M. Calahan, R. Young, G.T. Babcock and C.K. Chang, J. Am. Chem. Soc., 105, 634 (1983).

5. "Resonance Raman Detection of Fe-CO and Fe-C-O Bending Vibrations in Sterically Hindered Carbonmonoxy 'Strapped Hemes'. A Structural Probe of Fe-C-O Distortion", N.T. Yu, E.A. Kerr,B. Ward and C.K. Chang, Biochemistry, 22, 4534 (1983).

6. "Kinetic Study of CO and O2 Binding to Horse Heart Myoglobin Reconstituted with Synthetic Hemes Lacking Methyl and Vinyl Side Chains", C.K. Chang, B. Ward and S. Ebina, Arch. Biochem. Biophys., 231, 366 (1984).

7. "Spectral Properties of Protonated Schiff's Base Porphyrins and Chlorophylls. INDO-CI Calculations and Resonance Raman Studies", L.K. Hanson, P.M. Calahan, C.K. Chang, B. Ward and G.T. Babcock, J. Am. Chem. Soc., 106, 3950 (1984).

8. "Spectral Shifts Upon Reversible Modifications of CHO Peripheral Substituents in Porphyrin, Chlorin and Bacteriochlorin. A Phenomenological Explanation for the Red Shift of Protonated Schiff Bases", B. Ward, C.K. Chang and R. Young, J. Am. Chem. Soc., 106, 3943 (1984).

9. "Fine Tuning of Heme Reactivity: Hydrogen-Bonding and Dipole Interactions Affecting Ligand Binding to Hemes", C.K. Chang, B. Ward, R. Young and M.P. Kondylis, J. Macromol. Sci., Chem., A25(10-11), 1307-26.

10. "DNA Cleavage Specificity of a Group of Cationic Metalloporphyrins", B.Ward, A.S. Skorobogaty, J.C. Dabrowiak, Biochemistry, 25, 6875-6883 (1986).

11. "The DNA Binding Specificity of a Series of Cationic Metalloporphyrin Complexes", B. Ward, A.S. Skorobogaty and J.C. Dabrowiak, Biochemistry, 25, 7827 (1986).

12. "Molecular Recognition Between Oligopeptides and Nucleic Acids. Novel Imidazole Containing Oligopeptides Related to Netropsin Which Exhibit Altered DNA Sequence Specificity", J.W. Lown, K. Krowicki, U.G. Bhat, A.S. Skorobogaty, B. Ward and J.C. Dabrowiak, Biochemistry, 25, 7408-7416 (1986).

13. "Cationic Porphyrins as Probes of DNA Structure", S.D. Bromley, B. Ward and J.C. Dabrowiak, Nucleic Acids Res., 14, 9133 (1986).

14. "Quantitative Footprinting Analysis of the Netropsin-DNA Interaction", B. Ward, R. Rehfuss and J.C.Dabrowiak, J. Biomolecular Structure & Dynamics, 4, 685 (1987).

15. "DNA Binding Specificity of the Gold (III) Complex (C2H5)3PAuBr3", B. Ward and J.C. Dabrowiak, J. Am. Chem. Soc., 109, 3810 (1987).

16. "Determination of Netropsin-DNA Binding Constants from Footprinting Data", B. Ward, R. Rehfuss, J. Goodisman and J.C. Dabrowiak, Biochemistry, 27, 1198 (1988).

17. "Rate Enhancements in the DNase I Footprinting Experiment", B. Ward, R. Rehfuss, J. Goodisman and J.C. Dabrowiak, Nucleic Acids Res., 16, 1359 (1988).

18. "Interaction of Cationic Manganese Porphyrins with DNA. A Binding Model", G. Raner, B. Ward and J.C. Dabrowiak, J. Coord. Chem., 19, 17 (1988).

19. "Molecular Recognition Between Oligopeptides and Nucleic Acids - Rational Design of Sequence Specific DNA Binding Agents", K. Krowicki, M. Lee, J.A. Hartley, B. Ward, J.C. Dabrowiak and J.W. Lown, in Structure and Expression Vol. 2, DNA and its Drug Complexes, R.H. Sarma and M.H. Sarma eds., Adenine Press, 251

20. "Esperamicins, A Novel Class of Potent Antitumor Antibiotic: Mechanism of Action", B.H. Long, J. Golik, S. Forenza, B. Ward, R. Rehfuss, J.C. Dabrowiak, J. Catino, S. Musial, K.W. Brookshire, and T.W. Doyle, Proc. Natl. Acad. Sci. USA, 86, 2 (1989).

21. "Quantitative Footprinting Analysis Using a DNA-Cleaving Metalloporphyrin Complex", J.C. Dabrowiak, B. Ward and J. Goodisman, Biochemistry, 28, 3314 (1989).

22. "Quantitative Footprinting Analysis of the Actinomycin D-DNA Interaction", R. Rehfuss, B. Ward, J. Goodisman and J.C. Dabrowiak, in Molecular Basis of Specificity in Nucleic Acid-Drug Interactions, B. Pullman and J. Jortner eds., Kluwer Acad. Pub., Dordecht, Netherlands, 157-166 (1990)

23. "Site-Specific Binding Constants for Actinomycin-D on DNA Determined from Footprinting Data", J. Goodisman, R. Rehfus, B. Ward and J.C. Dabrowiak, Biochemistry, 31, 1046-1058 (1992)

24. "Type IIS Restriction Enzyme Footprinting I. Measurement of a Triplex Dissociation Constant with Eco57I at 25oC.", B. Ward, Nucleic Acids Research, 24, 2435-2440 (1996)

25. "A Comparison of Measured and Calculated Oligonucleotide Extinction Coefficients.", G. Kallanscrud, B. Ward, Analytical Biochemistry, 236, 134-138 (1996)

26. Quantitative DNA Footprinting." J.C. Dabrowiak, J. Goodisman and B. Ward, In Methods in Molecular Biology, Vol 90: Drug-DNA Interactions, K. Fox (ed.), Humana Press, Totowa. 23-42 (1997)

27. "Combinatorial Library Diversity: Probability Assessment of Library Populations." B. Ward and T. Juehne, Nucleic Acids Res. 26, 879-86 (1998)

28. "Mixed sequence triple helix binding: efficacy of natural, modified and surrogate bases in an AG motif third strand when opposed to pyrimidine interruptions in a polypurine target."F.M. Orson, J. Klysik, D.E. Bergstrom, B. Ward, G. A. Glass, P. Hua, and B.M. Kinsey, Nucleic Acids Res. 27, 810-816 (1999)

29. "A New Family of DNA cleavage Agents Directed by Triple-Helical Structures : Benzopyridoindole-EDTA conjugates."C. Marchand, J.S. Sun, B. Ward, C.H. Nguyen, E. Bisagni, T. Garestier and C. Hélène, in press.

30. "Exploring Ligan-DNA Space Using Type IIS Restriction Enzymes." Brian Ward, Combinatorial Chemistry & High Throughput Screening 5, 271-187 (2002).

31. "Slide Coating and DNA Immobilization Chemistries." Kathryn Aboytes, Jason Humphreys, Sonya Reis and Brian Ward In Microarrays for Beginners (working title), Eric Blalock (ed.), Kluwer and Associates, Norwell, manuscript submitted.

32. "Type IIS Restriction Enzyme Footprinting II. Quantitation of a Rapidly Equilibrating Ligand-DNA Complex." Brian Ward, manuscript in preparation.

33. "Type IIS Restriction Enzyme Footprinting III. Inhibition of Eco57I by a Sequence Specific Ligand is a Competitive Process.", B. Ward, (working title) manuscript in preparation.

In association with Brian Ward

1. "Solid phase-supported thymine dimers for the construction of dimer-containing DNA by combined chemical and enzymatic synthesis: a potentially general method for the efficient incorporation of modified nucleotides into DNA", Ordoukhanian P, Taylor JS, Nucleic Acids Res. 25, 3783-6 (1997)

Posters/Presentations

1. "Synthesis of an EDTA Phosphoramidite Suitable for Automated DNA Synthesis and Preliminary Strand Scission of ssM13mp18 with 5'FeEDTA-M13 Hybridization Probe", Brian Ward, Pharmaceutical Design: Nucleic Acid Binding Drugs, Palo Alto, CA, February 1994, poster.

2. "Type IIS Restriction Enzyme Footprinting. A Sigma Invention!", B. Ward, Sigma Chemical Co., St. Louis, December 1995, seminar.

3. "Type IIS Restriction Enzyme Footprinting: Measurement of a TFO Dissociation Constant", Brian Ward, Antisense Therapeutics, San Diego, CA, February 1996, poster.

4. "A Comparison of Calculated and Measured Oligonucleotide Extinction Coefficients", Gary Kallansrud and Brian Ward, Antisense Therapeutics, San Diego, CA, February 1996, poster.

5. "Type IIS Restriction Enzyme Footprinting 101: The Future's so Bright...", B. Ward, Bristol-Myers Squibb/ Sigma Chemical Company preferred customer agreement gala. Princeton, NJ December 1996, seminar.

Patents

1. "DNA's and Methods to Study Ligand-DNA Interactions Using Remote Cleaving Endonucleases", B. Ward, Status: Provisional Patent filed 12/96.

2. "Tracer Reagents that Enhance Reaction-Product Analysis", B. Ward, D. Ornitz, M. Deines, T. Bittick, Status: Patent Pending.

EXHIBIT C

OIPE
AUG 0 1 2002
PATENT & TRADEMARK OFFICE

Interview Summary

Application No.	Applicant(s)	
09/610,935	WARD ET AL.	
Examiner	**Art Unit**	
Bradley L. Sisson	1634	

All participants (applicant, applicant's representative, PTO personnel):

(1) *Bradley L. Sisson*.

(2) *Edward J. Hejlek*.

(3)*Brian Ward;* (5) *JoAnne Kerschner, Mgr. R&D*.

(4)*Bill Brizzard, Technology Transfer Mgr*.

Date of Interview: *20 February 2002* .

Type: a)☒ Telephonic b)☐ Video Conference
c)☐ Personal [copy given to: 1)☐ applicant 2)☐ applicant's representative]

Exhibit shown or demonstration conducted: d)☒ Yes e)☐ No.
If Yes, brief description: *Mr. Hejlek had faxed to Mr. Sisson a proposed amendment to the claims (attached)* .

Claim(s) discussed: *1-41* .

Identification of prior art discussed: _____ .

Agreement with respect to the claims f)☒ was reached. g)☐ was not reached. h)☐ N/A.

Substance of Interview including description of the general nature of what was agreed to if an agreement was reached, or any other comments: *See Continuation Sheet* .

(A fuller description, if necessary, and a copy of the amendments which the examiner agreed would render the claims allowable, if available, must be attached. Also, where no copy of the amendments that would render the claims allowable is available, a summary thereof must be attached.)

i)☐ It is not necessary for applicant to provide a separate record of the substance of the interview(if box is checked).

Unless the paragraph above has been checked, THE FORMAL WRITTEN REPLY TO THE LAST OFFICE ACTION MUST INCLUDE THE SUBSTANCE OF THE INTERVIEW. (See MPEP Section 713.04). If a reply to the last Office action has already been filed, APPLICANT IS GIVEN ONE MONTH FROM THIS INTERVIEW DATE TO FILE A STATEMENT OF THE SUBSTANCE OF THE INTERVIEW. See Summary of Record of Interview requirements on reverse side or on attached sheet.

RECEIVED
AUG 0 5 2002
TECH CENTER 1600/2900

B. L. Sisson
Examiner's signature, if required

Examiner Note: You must sign this form unless it is an Attachment to a signed Office action.

Summary of Record of Interview Requirements

Manual of Patent Examining Procedure (MPEP), Section 713.04, Substance of Interview Must be Made of Record
A complete written statement as to the substance of any face-to-face, video conference, or telephone interview with regard to an application must be made of record in the application whether or not an agreement with the examiner was reached at the interview.

Title 37 Code of Federal Regulations (CFR) § 1.133 Interviews
Paragraph (b)
In every instance where reconsideration is requested in view of an interview with an examiner, a complete written statement of the reasons presented at the interview as warranting favorable action must be filed by the applicant. An interview does not remove the necessity for reply to Office action as specified in §§ 1.111, 1.135. (35 U.S.C. 132)

37 CFR §1.2 Business to be transacted in writing.
All business with the Patent or Trademark Office should be transacted in writing. The personal attendance of applicants or their attorneys or agents at the Patent and Trademark Office is unnecessary. The action of the Patent and Trademark Office will be based exclusively on the written record in the Office. No attention will be paid to any alleged oral promise, stipulation, or understanding in relation to which there is disagreement or doubt.

The action of the Patent and Trademark Office cannot be based exclusively on the written record in the Office if that record is itself incomplete through the failure to record the substance of interviews.

It is the responsibility of the applicant or the attorney or agent to make the substance of an interview of record in the application file, unless the examiner indicates he or she will do so. It is the examiner's responsibility to see that such a record is made and to correct material inaccuracies which bear directly on the question of patentability.

Examiners must complete an Interview Summary Form for each interview held where a matter of substance has been discussed during the interview by checking the appropriate boxes and filling in the blanks. Discussions regarding only procedural matters, directed solely to restriction requirements for which interview recordation is otherwise provided for in Section 812.01 of the Manual of Patent Examining Procedure, or pointing out typographical errors or unreadable script in Office actions or the like, are excluded from the interview recordation procedures below. Where the substance of an interview is completely recorded in an Examiners Amendment, no separate Interview Summary Record is required.

The Interview Summary Form shall be given an appropriate Paper No., placed in the right hand portion of the file, and listed on the "Contents" section of the file wrapper. In a personal interview, a duplicate of the Form is given to the applicant (or attorney or agent) at the conclusion of the interview. In the case of a telephone or video-conference interview, the copy is mailed to the applicant's correspondence address either with or prior to the next official communication. If additional correspondence from the examiner is not likely before an allowance or if other circumstances dictate, the Form should be mailed promptly after the interview rather than with the next official communication.

The Form provides for recordation of the following information:
- Application Number (Series Code and Serial Number)
- Name of applicant
- Name of examiner
- Date of interview
- Type of interview (telephonic, video-conference, or personal)
- Name of participant(s) (applicant, attorney or agent, examiner, other PTO personnel, etc.)
- An indication whether or not an exhibit was shown or a demonstration conducted
- An identification of the specific prior art discussed
- An indication whether an agreement was reached and if so, a description of the general nature of the agreement (may be by attachment of a copy of amendments or claims agreed as being allowable). Note: Agreement as to allowability is tentative and does not restrict further action by the examiner to the contrary.
- The signature of the examiner who conducted the interview (if Form is not an attachment to a signed Office action)

It is desirable that the examiner orally remind the applicant of his or her obligation to record the substance of the interview of each case unless both applicant and examiner agree that the examiner will record same. Where the examiner agrees to record the substance of the interview, or when it is adequately recorded on the Form or in an attachment to the Form, the examiner should check the appropriate box at the bottom of the Form which informs the applicant that the submission of a separate record of the substance of the interview as a supplement to the Form is not required.

It should be noted, however, that the Interview Summary Form will not normally be considered a complete and proper recordation of the interview unless it includes, or is supplemented by the applicant or the examiner to include, all of the applicable items required below concerning the substance of the interview.

A complete and proper recordation of the substance of any interview should include at least the following applicable items:
1) A brief description of the nature of any exhibit shown or any demonstration conducted,
2) an identification of the claims discussed,
3) an identification of the specific prior art discussed,
4) an identification of the principal proposed amendments of a substantive nature discussed, unless these are already described on the Interview Summary Form completed by the Examiner,
5) a brief identification of the general thrust of the principal arguments presented to the examiner,
   (The identification of arguments need not be lengthy or elaborate. A verbatim or highly detailed description of the arguments is not required. The identification of the arguments is sufficient if the general nature or thrust of the principal arguments made to the examiner can be understood in the context of the application file. Of course, the applicant may desire to emphasize and fully describe those arguments which he or she feels were or might be persuasive to the examiner.)
6) a general indication of any other pertinent matters discussed, and
7) if appropriate, the general results or outcome of the interview unless already described in the Interview Summary Form completed by the examiner.

Examiners are expected to carefully review the applicant's record of the substance of an interview. If the record is not complete and accurate, the examiner will give the applicant an extendable one month time period to correct the record.

Examiner to Check for Accuracy

If the claims are allowable for other reasons of record, the examiner should send a letter setting forth the examiner's version of the statement attributed to him or her. If the record is complete and accurate, the examiner should place the indication, "Interview Record OK" on the paper recording the substance of the interview along with the date and the examiner's initials.

Continuation of Substance of Interview including description of the general nature of what was agreed to if an agreement was reached, or any other comments: In response to the proposed amendment of 2/20/02, Mr. Sisson expressed concern over the aspect of the subject application providing an adequate written description of the genus of compositions encompassed by the claims, noting that while Example 1 may set forth a procedure by which alternative embodiments may be produced, did not necessarily provide an adequate written description of he actual compositions per se, but rather identified possible starting materials that could be used in making certain embodiments. In response to argument from Mr. Hejlek that one could easily substitute one polymerase or dye for another, and in response to argument from co-inventor Mr. Ward, that polymerase activity would not likely be lost should such a substitution be performed, Mr. Sisson noted that such arguments may well be persuasive if the claims were drawn to a method of making the composition, and not to the actual product. Mr. Sisson noted that the level of difficulty or ease in producing the product was not dispositive as the rejection was not one of enablement, but rather, one of written description. Mr. Sisson indicated that the decision in Eli Lilly supported the position that one may be enabled for a method of making the product yet not satisfy the written description requirement for the product. And Mr. Sisson indicated that the decision in In re Shokal, 113 USPQ 283 (CCPA 1957) provided guidance as to the number of examples that need to be provided in order to adequately describe the members of a genus. Mr. Hejlek, upon discussion with Ms. Kerschner, Messrs. Ward and Brizzard suggested limiting the claims to a composition that comprises a thermostable DNA polymerase. Mr. Sisson noted that Taq, Pfu, Tth, Tfl, and Tli are commercially available thermostable DNA polymerases and would be encompassed by such a claim. Mr. Hejlek indicated that he would be submitting a declaration that shows how the exemplification of a composition comprising Taq polymerase constitutes and adequate written description of the genus of all thermostable DNA polymerases, and that this showing will be balanced against the guidance found in In re Shokal (number of species described in the specification v. the breadth of scope of the claims).
Mr. Hejlek indicated that he has corrected drawings ready for submission.

EXHIBIT A

Protein family review

Replicative DNA polymerases

M Mar Albà

Address: Wohl Virion Centre, Department of Immunology and Molecular Pathology, University College London, Cleveland Street, London, W1T 4JF, UK. E-mail: m.alba@ucl.ac.uk

Published: 12 January 2001

Genome Biology 2001, 2(1):reviews3002.1–3002.4

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2001/2/1/reviews/3002

© BioMed Central Ltd (Print ISSN 1465-6906; Online ISSN 1465-6914)

Summary

Replicative DNA polymerases are essential for the replication of the genomes of all living organisms. On the basis of sequence similarities they can be classified into three types. Type A polymerases are homologous to bacterial polymerases I, Type B comprises archaebacterial DNA polymerases and eukaryotic DNA polymerase α, and the bacterial polymerase III class make up type C. Structures have been solved for several type A and B polymerases, which share a similar architecture. The structure of type C is not yet known. The catalytic mechanism of all three types involves two metal-ion-binding acidic residues in the active site. Replicative polymerases are constitutively expressed, but their activity is regulated through the cell cycle and in response to different growth conditions.

Gene organization and evolutionary history

Classification

On the basis of sequence similarities, DNA polymerases can fall into three groups: type A, type B and type C, which have homology to *polA* (pol I), *polB* (pol II) and *polC* (pol III) from *Escherichia coli*, respectively [1,2]. Type C polymerases are not known to share structural similarity with types A and B, so they will not be covered in much detail in this review. In addition to replicative DNA polymerases, these groups also include polymerases involved in other types of DNA synthesis and in DNA repair. For example in bacteria the main replicative DNA polymerase is pol III (type C), while DNA polymerase I (type A) is not essential for replication and DNA polymerase II (type B) is only present in a few bacteria. Pol I has a role in nucleotide excision repair and in the processing of Okazaki fragments that are generated on the lagging strand during DNA replication, while pol II is known to be involved in repair of DNA cross-links. Polymerase delta from eukaryotes belongs to the same polymerase group as the replicative DNA polymerase α (type B). The role of polymerase delta in replication however is not clear although there is evidence that it participates in post-replicational DNA repair.

Replicative DNA polymerases from some bacteriophages (T3, T5 and T7) and eukaryotic mitochondrial DNA polymerases have homology to bacterial polymerases I and are therefore type A polymerases. Eukaryote replicative polymerase α, archaebacterial DNA polymerases, viral DNA polymerases, DNA polymerases encoded in mitochondrial plasmids of various fungi and plants and some bacteriophage polymerases (T4 and RB69) all belong to type B. The bacterial DNA polymerase III class, members of which are responsible for most of the replicative DNA synthesis in bacteria, are type C DNA polymerases [1]. The three types share no obvious sequence similarity, but types A and B are structurally similar to each other (see below). In general, a single gene for each type is found in the different organisms, but there are exceptions; for example, some bacteria have several genes for type C DNA polymerases. Little is known about the structure of the DNA polymerase α gene in eukaryotes, although the mouse gene is known to contain four exons.

Characteristic structural features

Sequence features

Type A polymerases contain three conserved motifs: A, B (Prosite signature PS00447 [3]) and C. Motifs A and C (Figure 1) are part of the catalytic site, whereas motif B is involved in the binding of dNTPs. In type B polymerases, up to six regions of sequence homology have been identified. Regions I (Prosite signature PS00116) and II form part of the active site and are considered to be equivalent to polymerase

type A motifs C and A, respectively (Figure 1). In these regions, metal-ion-binding aspartic acid residues are in equivalent structural positions [4,5].

In addition to the catalytic domain, DNA polymerases often have additional domains required for editing activity, excision of Okazaki primers during replication (structure-specific 5' nuclease activity), or for interactions with other proteins. Bacterial and archaebacterial replicative DNA polymerases can contain a 3'-to-5' proofreading exonuclease domain, and the eukaryotic α polymerase catalytic subunit contains a zinc finger domain required for interactions with other subunits of the polymerase-primase complex [6]. A 35-residue fragment with the potential to form a Zn-finger has also been identified in type C DNA polymerase from bacteria. This region appears essential for the proper formation and/or function of the enzyme's polymerase site [7].

Structural features

The *E. coli* DNA polymerase I structure was detemined in 1985 [8], and since then, several other type A DNA polymerase structures, such as T7 DNA polymerase [9] (Figure 2a), have been solved. For type B DNA polymerases three crystal structures are available, from phage RB69 [10] and from two *Thermococcus* species of archaebacteria (Figure 2b) [11,12].

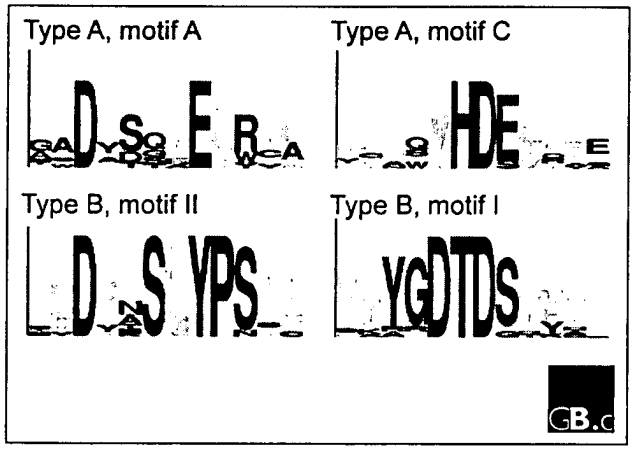


Figure 1
Logo representation of DNA polymerase domains from representative polymerase sequences. The sequences were extracted from SwissProt links in Prosite [3] to the DNA polymerase type A (PS00447) and DNA polymerase type B (PS00116) families. The families were initially found using 'DNA polymerase' as keyword. Sequence logos show the relative representation of the different amino acids at each sequence position in terms of the size of the appropriate letter in the single-letter amino acid code; the largest letters are the most conserved. They were constructed using the online server WebLogo [17]. Type A DNA polymerases (10 sequences) were from *E. coli* (polA), *Bacillus subtilis*, *Thermus aquaticus*, *Aquifex aeolicus*, *Mycobacterium tuberculosis*, phage T3, phage T5, phage T7, mitochondrial *Saccharomyces cerevisiae* DNA polymerase and mitochondrial *Homo sapiens* DNA polymerases. Type B DNA polymerases (15 sequences) were from *H. sapiens* (α and δ), *S. cerevisiae* (α and δ), *Drosophila melanogaster* (α), phage T4, phage RB69, *Chlorella* virus NY-2A, *Zea mays* plasmid S-1, herpes simplex virus 1, vaccinia virus, African swine fever virus, *Archaeoglobus fulgidus*, *Pyrococcus horikoshii*, *E. coli* (polB).

The catalytic domains of type A and B DNA polymerases have a common overall architecture, which resembles a right hand and consists of 'thumb', 'palm' and 'fingers' domains [4] (Figure 2a,b). A similar structure and catalytic mechanism is shared by other families of polymerases, such as eukaryotic DNA polymerase β, reverse transcriptases and RNA-dependent RNA polymerases. The most conserved region is the palm domain, which contains the catalytic site. The fingers and thumb have somewhat different arrangements in the two families, although the thumb always contains parallel or anti-parallel α helices that appear to interact with the minor groove of the primer-template complex, and the fingers have an α helix with conserved sidechains positioned at the blunt end of the primer-template complex.

Localization and function

The replicative polymerases are required for the faithful replication of the genetic material and they perform this role by attaching the appropiate nucleotide to the nascent strand to match the template strand. The replication machinery, which includes a number of other well-conserved enzymes such as helicases and primases, assembles at the replication origin, where the replicative DNA polymerase initiates DNA synthesis using short DNA or RNA primers. All known DNA polymerases synthesize DNA in a 5'-to-3' direction.

The replication of DNA occurs before cell division. Replicative DNA polymerase genes are housekeeping genes and in eukaryotes the enzyme is found in all nuclei. The polymerase activity can be controlled in a cell-cycle-dependent manner or in response to different environmental conditions, however. For example, in mammals the activity of replicative DNA polymerase α is regulated by phosphorylation during the cell cycle [13] and its expression is stimulated by specific transcription factors during growth [14].

Enzyme mechanism

DNA synthesis is mediated by transfer of a phosphoryl group from the incoming dNTP to the DNA 3' OH, liberating a pyrophosphate and forming a new DNA phosphodiester bond. This reaction is catalysed by a mechanism that involves two metal ions, normally Mg^{2+}, with the participation of two aspartic acid residues that are structurally conserved among the different enzymes [5,6]. These are the aspartic acids conserved in motifs A and II, and C and I, from type A and type B polymerases, respectively (Figure 1). The first metal ion activates the 3'-OH for attack of the α phosphate of the incoming dNTP and the second metal ion

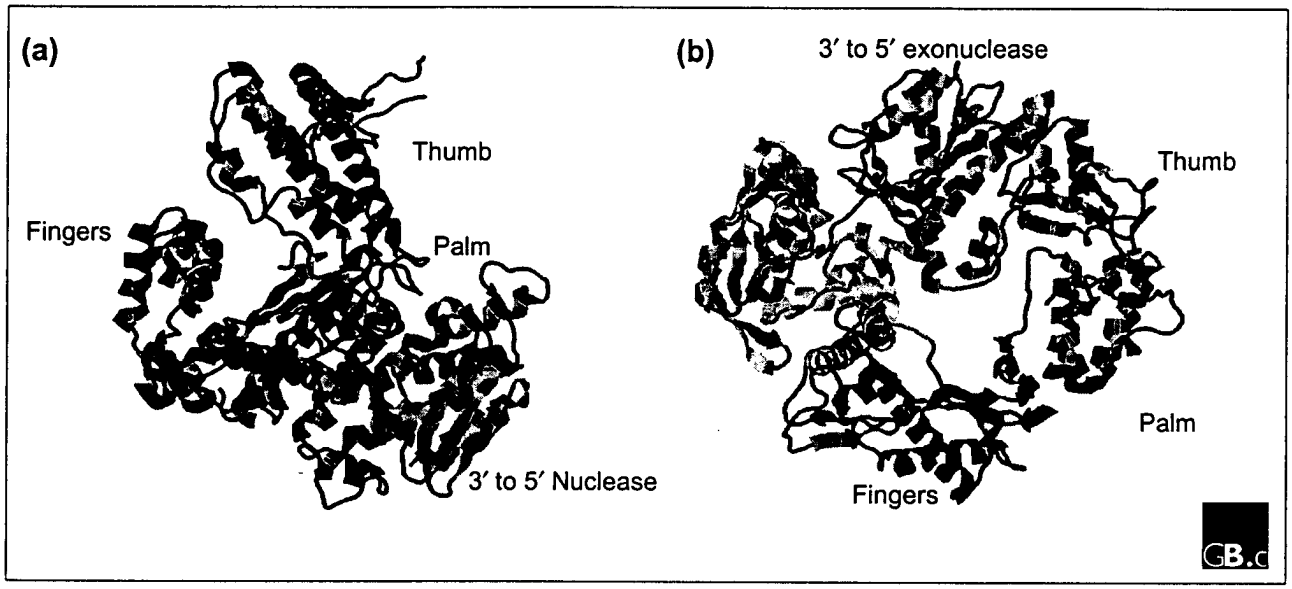


Figure 2
DNA polymerase protein structures. The figures were generated with the free software Rasmol (RasWin Molecular Graphics, Windows Version 2.4) and files extracted from the Protein Data Bank (PDB) [18]. **(a)** Bacteriophage T7 DNA polymerase (type A; PDB code: 1t7p) [9]. Blue, catalytic domain; green, 3'-to-5' exonuclease domain. **(b)** *Thermococcus gorgorianus* type B DNA polymerase [11] (PDB code: 1tgo). Blue, catalytic domain; green, 3'-to-5' exonuclease domain; orange, amino-terminal domain; yellow, linker region.

stabilizes the negative charge that builds up on the leaving oxygen and chelates the β- and γ-phosphates.

The catalytic mechanism of the type C DNA polymerase catalytic subunit may be similar to type A and B polymerases. Three conserved aspartic acid residues have been identified that strongly affect the polymerase activity and that may therefore be involved in the coordination of metal ions in the active site [15].

Frontiers

Replication in prokaryotes and eukaryotes is a relatively well studied pathway and a number of solved type A and type B DNA polymerase structures have provided evidence of a common architecture and active-site mechanism. Similar features have been observed in other non-replicative DNA polymerases. The determination of structures of type C DNA polymerases will be interesting, as it may show whether all polymerases can be unified as a single structural class.

An interesting aspect of replicative DNA polymerases is the diversity of enzymes, covering at least three different sequence families, that perform the same function. DNA replication is an ancestral function and as such is expected to be well conserved. The presence of different types of functionally homologous polymerases in different organisms suggests a complex evolution, however. One hypothesis, which involves transfer of DNA polymerase genes from viruses to eukaryotic genomes, has recently been put forward to attempt to address this question [16]. There is clearly much still to be learnt about these important enzymes.

References

1. Delarue M, Poch O, Tordo N, Moras D, Argos P: **An attempt to unify the structure of polymerases.** *Prot Engineering* 1990, **3**:461-467.
Pioneer work in the identification of conserved regions in various polymerase types and comparison with the *E. coli* polymerase I fold.
2. Braithwaite DK, Ito J: **Compilation, alignment, and phylogenetic relationships of DNA polymerases.** *Nucleic Acids Res* 1993, **21**:787-802.
A collection of aligned type A, B and C DNA polymerase sequences. Although somewhat outdated, this article is still a good reference for DNA polymerase families.
3. **Prosite** [http://www.expasy.ch/prosite/]
Database of protein domains.
4. Joyce CM: **Choosing the right sugar: how polymerases select a nucleotide substrate.** *Proc Natl Acad Sci USA* 1997, **94**:1619-1622.
Comparison of residues in the active site of different types of DNA and RNA polymerases. Emphasis is put on the possible substrate discriminative mechanisms.
5. Steitz TA: **DNA polymerases: structural diversity and common mechanisms.** *J Biol Chem* 1999, **274**:17395-17398.
An excellent review on the structure of DNA polymerases.
6. Mizuno T, Yamagishi K, Miyazawa H, Hanaoka F: **Molecular architecture of the mouse DNA polymerase alpha-primase complex.** *Mol Cell Biol*, 1999 **19**:7886-7896.
A detailed study of interactions between the four different subunits of the polymerase complex by selectively expressing combinations of them in cultured mammalian cells.
7. Barnes MH, Leo CJ, Brown NC: **DNA polymerase III of Gram-positive eubacteria is a zinc metalloprotein conserving an essential finger-like domain.** *Biochemistry* 1998, **37**:15254-15260.
Analysis of the functionality of the finger-like domain fragment in pol III.

8. Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA: **Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP.** *Nature* 1985, **313**:762-766.
Pioneer work in the determination of a DNA polymerase structure.
9. Doublié S, Tabor S, Long AM, Richardson CC, Ellenberger T: **Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution.** *Nature* 1998, **391**:251-258.
Includes the complete structure of the bacteriophage T7 DNA polymerase (type A), with polymerase and 3' to 5' exonuclease domains.
10. Wang J, Sattar AKMA, Wang CC, Karam JD, Konigsberg WH, Steitz TA: **Crystal structure of a pol α replication DNA polymerase from bacteriophage RB69.** *Cell* 1997, **89**:1087-1089.
First structure of a type B DNA polymerase to be determined.
11. Hopfner KP, Eichinger A, Engh RA, Laue F, Ankenbauer W, Huber R, Angerer B: **Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*.** *Proc Natl Acad Sci USA* 1999, **96**:3600-3605.
Structure of an archaebacterial DNA polymerase and comparison to the RB69 DNA polymerase structure for the identification of features that confer increased thermostability to Taq polymerase.
12. Rodriguez AC, Park HW, Mao C, Beese LS: **Crystal structure of a pol alpha family DNA polymerase from the hyperthermophilic archaeon thermococcus sp. 9 degrees N-7.** *J Mol Biol* 2000, **299**:471-487.
Crystal structure of another thermostable archaeal DNA polymerase. A putative RNA-binding domain is identified in the amino-terminal region.
13. Voitenleitner C, Rehfuess C, Hilmes M, O'Rear L, Liao PC, Gage DA, Ott R, Nasheuer HP, Fanning E: **Cell cycle-dependent regulation of human DNA polymerase alpha-primase activity by phosphorylation.** *Mol Cell Biol* 1999, **19**:646-656.
An analysis of the phosphorylation of the polymerase α catalytic and p68 subunits by cyclin A/cdk2 and cyclin E/cdk2 kinases and its effect on initiation of replication.
14. Izumi M, Yokoi M, Nishikawa NS, Miyazawa H, Sugino A, Yamagishi M, Yamaguchi M, Matsukage A, Yatahaia F, Hanaoka F: **Transcription of the catalytic 180-kDa subunit gene of mouse DNA polymerase alpha is controlled by E2F, an Ets-related transcription factor, and Sp1.** *Biochim Biophys Acta* 2000, **1492**:341-352.
Identification of an E2F-binding site near the transcription start site that is essential for growth-dependent stimulation of gene expression.
15. Pritchard AE, McHenry CS: **Identification of the acidic residues in the active site of DNA polymerase III.** *J Mol Biol* 1999, **285**:1067-1080.
Comparison of sequences and mutagenesis analysis were used to identify residues in the active site of the catalytic subunit of DNA polymerase III.
16. Villarreal LP, DeFilippis VR: **A hypothesis for DNA viruses as the origin of eukaryotic replication proteins.** *J Virol* 2000, **74**:7079-7084.
Sequence comparison and phylogenetic trees of type A DNA polymerases from a variety of organisms, including DNA viruses. The hypothesis is formulated that replication genes, including those encoding DNA polymerases, were transferred from viruses to eukaryotes during evolution.
17. **WebLogo** [http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi]
A service for constructing logos (as in Figure 1) from sequence alignments.
18. **Protein Data Bank (PDB)** [http://pdb.ccdc.cam.ac.uk/pdb/]
Database of protein structures.

LIFE
...n an RNA probe kit
see inside back cover

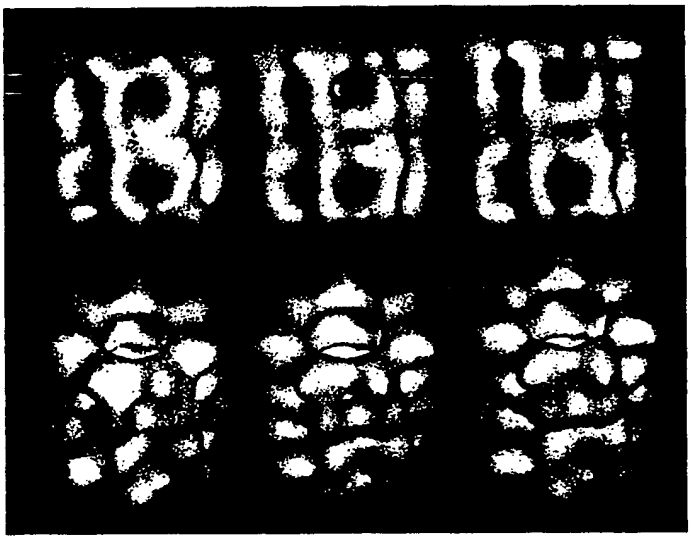
Structural Biology

Protein–nucleic acid interactions
Edited by Timothy J Richmond and Thomas A Steitz

Folding and binding
Edited by Eugene Shakhnovich and Alan R Fersht

Paper alert

Web alert



This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes

Chad A Brautigam* and Thomas A Steitz*†

New levels in the understanding of DNA replication have been achieved from recent crystal structure determinations of several DNA polymerases and their substrate complexes. The structure of an α family DNA polymerase from bacteriophage RB69 shows some similarities, but also considerable differences in structure and organization from the pol I family DNA polymerases. Also, the functions of three polymerase domains and their conserved residues have been clarified by studying structures of pol I family DNA polymerases complexed to their substrates. These structures also confirm that an identical two-metal-ion catalytic mechanism proposed previously is used by both the nonhomologous pol I and pol β family DNA polymerases.

Addresses
*Department of Molecular Biophysics and Biochemistry, †Department of Chemistry, *Howard Hughes Medical Institute, Yale University, PO Box 208114, New Haven, CT 06520-8114, USA

Current Opinion in Structural Biology 1998, 8:54–63

http://biomednet.com/elecref/0959440X00800054

© Current Biology Ltd ISSN 0959-440X

Abbreviations

ddNTP	dideoxynucleoside triphosphate
gp43	gene 43 protein
KF	Klenow fragment
pol I	DNA polymerase I
pol α	DNA polymerase α
pol β	DNA polymerase β
RT	reverse transcriptase
T7 pol	T7 DNA polymerase
***Taq* pol**	*Taq* DNA polymerase

Introduction

Significant advances in the structural biology of DNA polymerases have been made on two fronts. First, the library of known polymerase structures has expanded to include the gene 43 protein (gp43) from bacteriophage RB69 [1••], a member of the human DNA polymerase α family, offering insights into eukaryotic replicative polymerases. This expanded library also includes structures of polymerases from bacteriophage T7 [2••] and a *Bacillus stearothermophilus* strain [3•]—examples of polymerases homologous to DNA polymerase I from *Escherichia coli*. On the second front, a combination of structural and biochemical techniques has expanded our understanding of several properties of DNA polymerases, such as substrate binding [4••,5•,6••], processivity [2••,6••], fidelity [2••,7•], and nucleotidyl transfer [2••]. Although DNA polymerases often occur as multifunctional polypeptides, we shall restrict the scope of this review to the polymerase domains.

The structure of gp43 from bacteriophage RB69

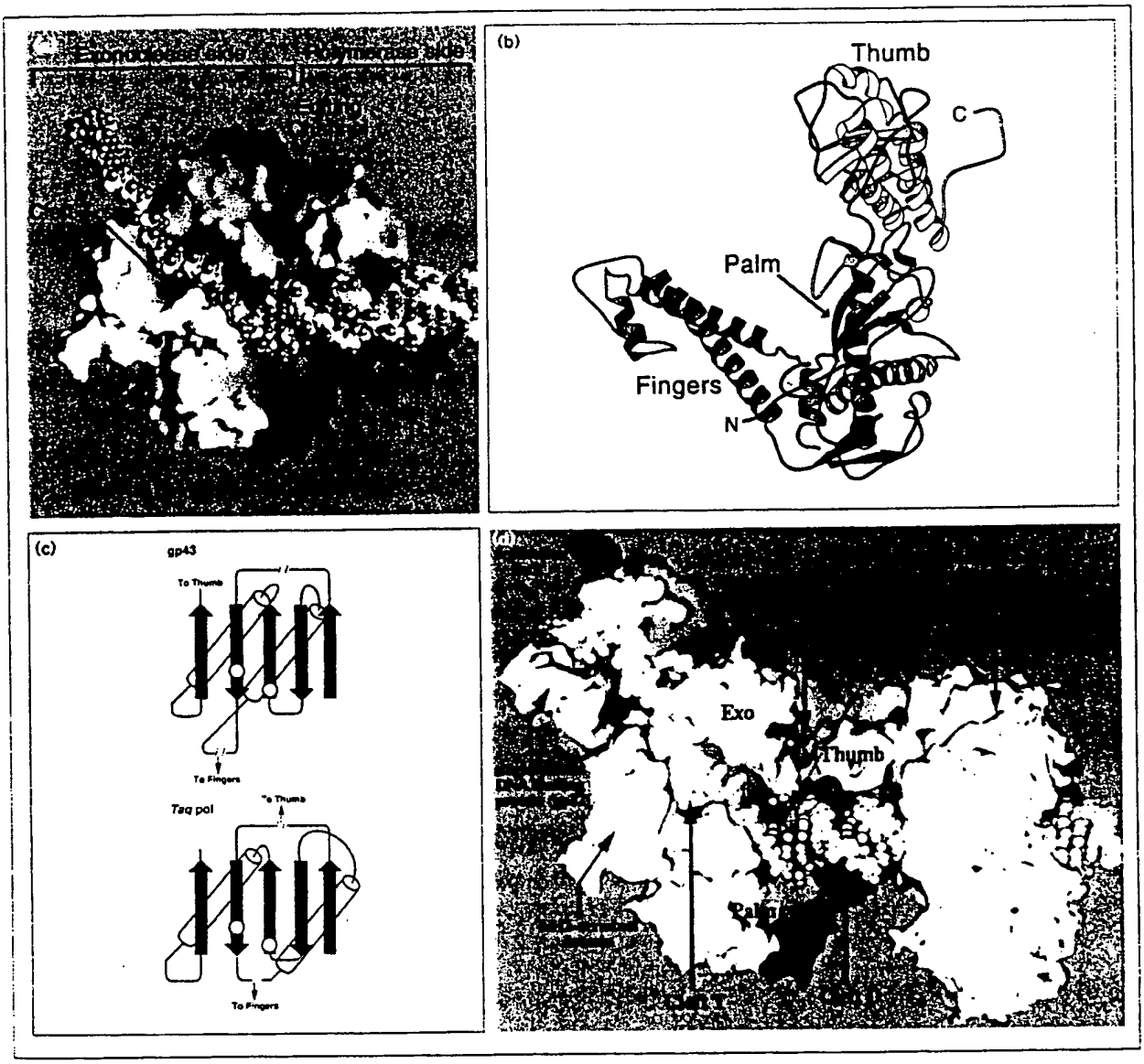
The gp43 protein from bacteriophage RB69, which is a close homologue of bacteriophage T4, is primarily responsible for the replication of the viral genome [8]. It shares both sequence and functional homology with human DNA polymerase α (pol α), an enzyme implicated the replication of chromosomal DNA. The gp43 protein contains the six conserved sequence hallmarks, known as regions I–VI, that are common to all pol α-like or 'B family' polymerases [9]. Functionally, it shares several attributes with these polymerases: it is responsible for both leading- and lagging-strand synthesis, it is tethered to the substrate DNA by a protein processivity factor and it is stimulated by a single-stranded DNA binding protein [8]. Thus, knowledge gained from the study of gp43 will be applicable to eukaryotic replicative DNA polymerases.

The crystal structure of RB69 gp43 at 2.8 Å resolution [1••] reveals some surprising features along with some known structural motifs. The overall structure of the 103-kDa polypeptide is that of a disc with a small hole in the center (Figure 1). Arrayed around this central cavity are five structurally and functionally distinct domains that form three clefts, termed cleft D, cleft T, and the editing channel (Figure 1a). Approximately one-half of the disc is formed by the polymerase portion of gp43, which adopts the right hand shape common to all known polymerase structures. It comprises three domains: the fingers, palm and thumb (Figure 1b). The catalytic center of the polymerase is located on the palm domain, near the central cavity. The disc is completed by an N-terminal domain and the 3′–5′ exonuclease domain, which is homologous to the Klenow fragment (KF) editing domain. The structure of this half of the disc is nearly identical to that of the homologous KF domain from bacteriophage T4 gp43 determined by Wang *et al.* [10].

One unexpected feature of the RB69 gp43 structure is the position of the proof-reading domain with respect to the polymerase domain, which is opposite to their arrangement in the DNA polymerase I (pol I) family polymerases. Also, unlike the pol I family, a single-stranded tetranucleotide binds in the editing channel so that no translation of the primer–template duplex is required for the primer terminus to bind either in this channel or at the polymerase active site. In *Taq* DNA polymerase (*Taq* pol), the primer–template duplex bound to the polymerase active site is translocated relative to the frayed duplex DNA bound with its primer terminus in the KF-like proof-reading site [4••,11]. As is generally the case for the pol α family, the editing activity of this phage polymerase

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Figure 1



The structure of gp43, the α family DNA polymerase from phage RB69. **(a)** A surface representation of the protein with homology modeled DNA. The polymerase half of the enzyme is on the right, whereas the exonuclease portion is on the left. The three clefts are labeled. The DNA in the structure has been modeled [1••] by comparison with the *Taq* pol protein–DNA co-crystal structure. The duplex product fits into cleft D while the modeled 5′ template was placed in cleft T. The central cavity is obscured by the DNA model. **(b)** A ribbon representation of the polymerase domain of gp43. The structure is color-coded by domain, with the fingers in the darkest gray, the thumb in the lightest gray, and the palm in an intermediate shade. The palm is dominated by an antiparallel β sheet. The thumb is mainly α-helical, but also contains some β strands. The unusual fingers feature two very long α helices. **(c)** Schematic representation showing the similar overall topologies of the palm domains of gp43 and *Taq* pol. Both comprise an antiparallel β sheet flanked on the noncatalytic side by two α helices. The positions of the conserved catalytic aspartates are shown as open circles. **(d)** A model of DNA and accessory proteins bound to gp43 [1••]. The DNA is modeled as in (a). The clamp protein, gp45, is proposed to interact with both the DNA and gp43 at the exit point of cleft D. The single-stranded DNA-binding protein, gp32, is proposed to interact with the single-stranded template after it leaves cleft T. Reproduced with permission from [1••].

is about 1000-fold higher than the exonucleolytic activity of the *E. coli* pol I enzyme [12–16]. The structure of a single-stranded DNA substrate bound to the exonuclease active site of gp43 shows the ε-amino group of a lysine residue in close proximity to the DNA scissile phosphate [1••,10]. This lysine is conserved among the pol α family, but is not present in the pol I family. The ε-amino group of this residue may enhance the exonuclease activity of gp43

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

by stabilizing the pentacovalent transition state through an interaction with a negatively charged, non-bridging phosphate oxygen.

A comparison of the polymerase motif of gp43 with the analogous motif in a binary complex between *Taq* pol [4••] and double-stranded DNA shows that the topology of the palm domains is identical and thus, by homology modeling, suggests the direction of DNA replication when bound to gp43. The structures of the β-sheet portions of the gp43 and *Taq* pol palm domains are very similar (Figure 1c). By superimposing the homologous palm domains of the binary complex of *Taq* pol and gp43 and by assuming that the two polymerases bind DNA similarly, a model for primer–template bound to gp43 was built [1••]. It positions the double-stranded portion of the DNA substrate at cleft D (Figure 1d). Furthermore, cleft T is in a suitable position to bind the single-stranded DNA template. This model of DNA bound to these clefts allowed Wang *et al.* [1••] to surmise the positions of the accessory proteins of gp43 within a functioning replisome (Figure 1d). The processivity factor of RB69, gp45, like that of the T4 phage, is a sliding-clamp protein whose donut-like structure is homologous to that of the β subunit of DNA polymerase III from *E. coli* (J Kuriyan, personal communication; see also [17]). It is proposed that gp45 interacts with the long C-terminal tail of gp43 as well as the double-stranded DNA on the polymerase side of the disc, at the exit point of cleft D [1••]. This model is supported by the fact that removal of the C-terminal tail of gp43 eliminates gp45-generated processivity [18]. The single-stranded DNA-binding protein from RB69, gp32, probably associates with gp43 on the exonuclease side of the disc, since cleft T exits the polymerase there. Thus, taken in the context of other polymerase structures, the structure of gp43 allows a reasonable model of a portion of the eukaryotic replication apparatus to be inferred.

The structure of gp43 also clarifies some aspects of the interrelatedness of pol α family polymerases and the relationships between this family and other polymerase families [1••]. Using the gp43 structure, the sequence alignments of all pol α family polymerases can be extended beyond regions I–VI to a span of nearly 200 residues encompassing all six of the previously identified sequence hallmarks (Figure 2a). An earlier attempt [19] to align the sequences of all DNA and RNA polymerases resulted in the conclusion that the active-site motif DTD (residues 621–623) of pol α was equivalent to a DD motif (residues 184–185) in HIV-1 reverse transcriptase (RT) or a DE motif (residues 882–883) in KF. The threonine of the pol α sequence was aligned as an insertion between the two conserved acidic residues. The structures of gp43, KF and HIV-1 RT [20], however, show that these two acidic residues are in similar positions only in KF and HIV-1 RT. Across all polymerases, only the second D of the DTD motif and the first D of the DD(E) motif are structurally conserved [1••] (Figure 2b).

Other polymerase structures

The crystal structures of two homologues of DNA polymerase I from *E. coli* have been solved recently. The first, a Klenow-like fragment of the DNA polymerase I of a thermophilic *Bacillus* species [3•], has been solved at 2.1 Å resolution. The structure of this enzyme is very similar to Klenow fragment [21], especially in its polymerase domain. The second structure, a quaternary complex of T7 DNA polymerase (T7 pol), thioredoxin (a processivity factor), DNA, and incoming dideoxynucleoside triphosphate was solved to 2.2 Å [2••]. The complex shows that the three-dimensional structure of T7 DNA polymerase is also analogous to that of KF.

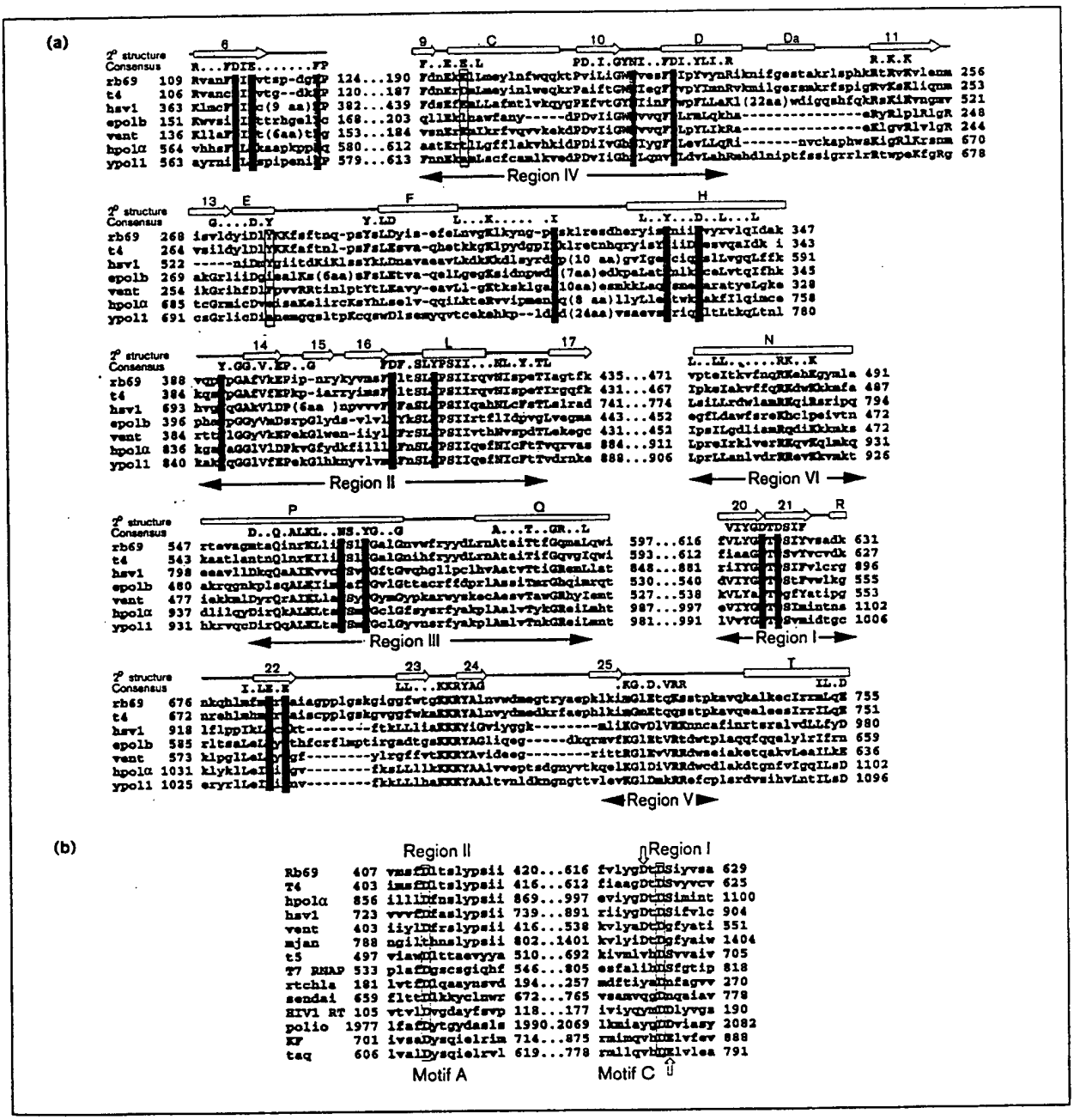
Structural comparisons among DNA polymerases

A comparison of all DNA polymerase structures elucidated to date reveals some interesting commonalities and some striking differences. Presently, the structures of enzymes from four DNA polymerase families [9] have been established: RB69 gp43 from the pol α or 'B family' [1••]; KF [21], *Taq* pol [22], T7 pol [2••], and both the *Bacillus* and *Thermus aquaticus* KF analogues [3•,23] from the pol I family; rat and human DNA polymerase β (pol β) [5•,24,25] from the terminal transferase family [26]; and HIV-1 RT [20,27–29] and a fragment of Moloney murine leukemia virus RT [30] from the reverse transcriptase family. The palm domains of all DNA polymerases consist of a four- to six-stranded β sheet that is flanked on one side by two α helices. In the pol I, pol α and reverse transcriptase families, the topologies of the palm domain are the same. Even though the palm domain of pol β is mostly β sheet, its topology differs from that of the other polymerase families and thus was not derived from a common evolutionary ancestor. Beyond the palm domains, structural comparisons break down. The structures of the fingers domains vary widely from family to family, and, although the thumb domains are mainly α-helical, the detailed structures of these domains are also not related. Perhaps surprisingly, the fingers and thumb domains in all four families have arisen from different ancestors.

Although there are structural differences among the families of DNA polymerases, complexes of these enzymes with primer–template DNA show that the anatomical features of these enzymes play analogous roles across family boundaries, often using similarly oriented secondary structural elements. This can be illustrated by viewing a binary complex of blunt-ended double-stranded DNA and *Taq* pol, an enzyme whose polymerase motif is a close homologue of the polymerase motif of KF (Figure 3a) [4••,22]. The primer terminus of the DNA abuts the fingers domain of *Taq* pol and several acidic residues from the palm domain that are responsible for binding the catalytically essential metal ions are located close to the primer. The DNA leads away from the fingers and has several contacts with the palm domain. It then encounters the thumb domain, which has extensive contacts with

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Figure 2



Sequence comparisons among DNA polymerases. **(a)** A comparison within the pol α family. Included in this comparison are polymerases from RB69 and T4 (rb69 and t4), herpes simplex virus 1 (hsv1), *E. coli* (pol II, [epolb]), *Pyrococcus furiosis* (vent), humans (pol α, [hpolα]), and yeast (pol I, [ypol1]). Arrows and cylinders above the alignment denote secondary structure, and regions I–VI are shown below. Capital letters are consensus residues. A consensus sequence is shown above the alignment. Residues shown in the darkest shade are conserved carboxylates, medium gray residues are solvent-exposed residues that may be involved in substrate binding, and lightest gray are positively charged residues. Boxed residues are carboxylates whose roles are unclear from the gp43 structure. **(b)** Sequence conservation among all DNA polymerases. In addition to some of the polymerases enumerated in (a), those from phage T5 (t5), *Methanococcus jannaschii* (mjan), mitochondrial *Chlamydomonas rheinhardtii* DNA intron (rtchla), sendai virus L (sendai), HIV-1 (HIV-1 RT), polio virus (polio), *E. coli* (KF), *Thermus aquaticus* (taq), and T7 RNA polymerase (T7 RNAP) are shown. Region II of the pol α family polymerases corresponds to motif A of [19], while region I corresponds to motif B. This sequence alignment is based on the conserved positions of carboxylates in the palm domain. Only two carboxylates are completely conserved across all of these polymerases (shown in boxes). The first D of the DTD motif and the second carboxylate of the DD(E) motif (arrows) are not structurally equivalent. Reproduced with permission from [1••].

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

the DNA across its minor groove. Whereas the above statements specifically concern *Taq* pol, they also hold true for T7 pol ([2••]; Figure 4a), HIV-1 RT ([27]; Figure 3b) and also for the homology model-built DNA complex with RB69 gp43 ([1••]; Figure 3c). The ternary complex of pol β with DNA and incoming dideoxynucleoside triphosphate (ddNTP) is shown in Figure 3d. Here, the primer–template DNA abuts domain D, and domain B contacts the DNA across the minor groove. Based on this functional analogy, it is clear that domains D and B would most usefully be called fingers and thumb, respectively. We shall follow this nomenclature throughout this review, even though a previous inappropriate alignment of pol β with non-homologous polymerases has led to an opposite, functionally less useful naming of these domains [25]. Viewed with this assignment in mind, all of the statements made above regarding *Taq* pol are true for pol β and an additional similarity among the pol α, pol I, and terminal transferase families becomes evident. They all have an α helix located in the fingers domain near the primer terminus that houses residues responsible for nucleotide binding. No analogous α helix is present in the reverse transcriptases, but rather there is an anti-parallel β hairpin located at the primer terminus.

The enzyme–substrate complexes of *Taq* pol [4••] and T7 pol [2••], along with the changed assignment of pol β, allow us to resolve a controversy that exists in the field [31–33]. A structure of an editing complex between KF and DNA shows the double-stranded portion of the substrate bound between the thumb and the 3′–5′ exonuclease domains [11]. This led Steitz and colleagues to propose that when the DNA was bound at the polymerase active site, it occupies roughly the same position as in *Taq* pol (Figure 3a). However, the alignment of KF and pol β resulting from the previous inappropriate superposition of their non-homologous palm regions spawned a proposal that the DNA should be bound to KF in the opposite orientation [25]. The structures of *Taq* pol and T7 pol demonstrate unequivocally that the proposal of Steitz and coworkers is correct and that the latter proposal should be disregarded.

Structural and functional insights into the activities of DNA polymerases

The quaternary complex of T7 pol [2••], the structure of *Taq* pol bound to DNA [4••] and several new structures of pol β [5•,7•] are remarkable for the insight they give into the processivity, fidelity, catalytic mechanism and conformational flexibility of DNA polymerases.

Processivity

DNA polymerases have the unique feature that they may remain associated with the substrate primer–template over many catalytic cycles, that is, they are processive [8]. Replicative DNA polymerases need the aid of a protein factor in order to remain processive over the many thousands of nucleotide additions required for genome replication. T7 pol, a viral replicative DNA polymerase, commandeers the *E. coli* protein thioredoxin for this purpose [34]. The crystal structure of the quaternary complex of T7 pol and its substrates includes thioredoxin ([2••]; Figure 4a). The factor binds to a loop that protrudes from the tip of the polymerase thumb. Although electron density for DNA does not appear near thioredoxin due to disorder, extrapolation of the visible DNA helix positions the primer–template for a potential interaction with the factor (Figure 4a). Since all kinetic schemes for DNA polymerases include a step in the catalytic cycle during which the DNA is susceptible to dissociation from the enzyme [13,16,35,36], thioredoxin or other processivity factors act at this stage, either sterically or electrostatically hindering the dissociation of the DNA.

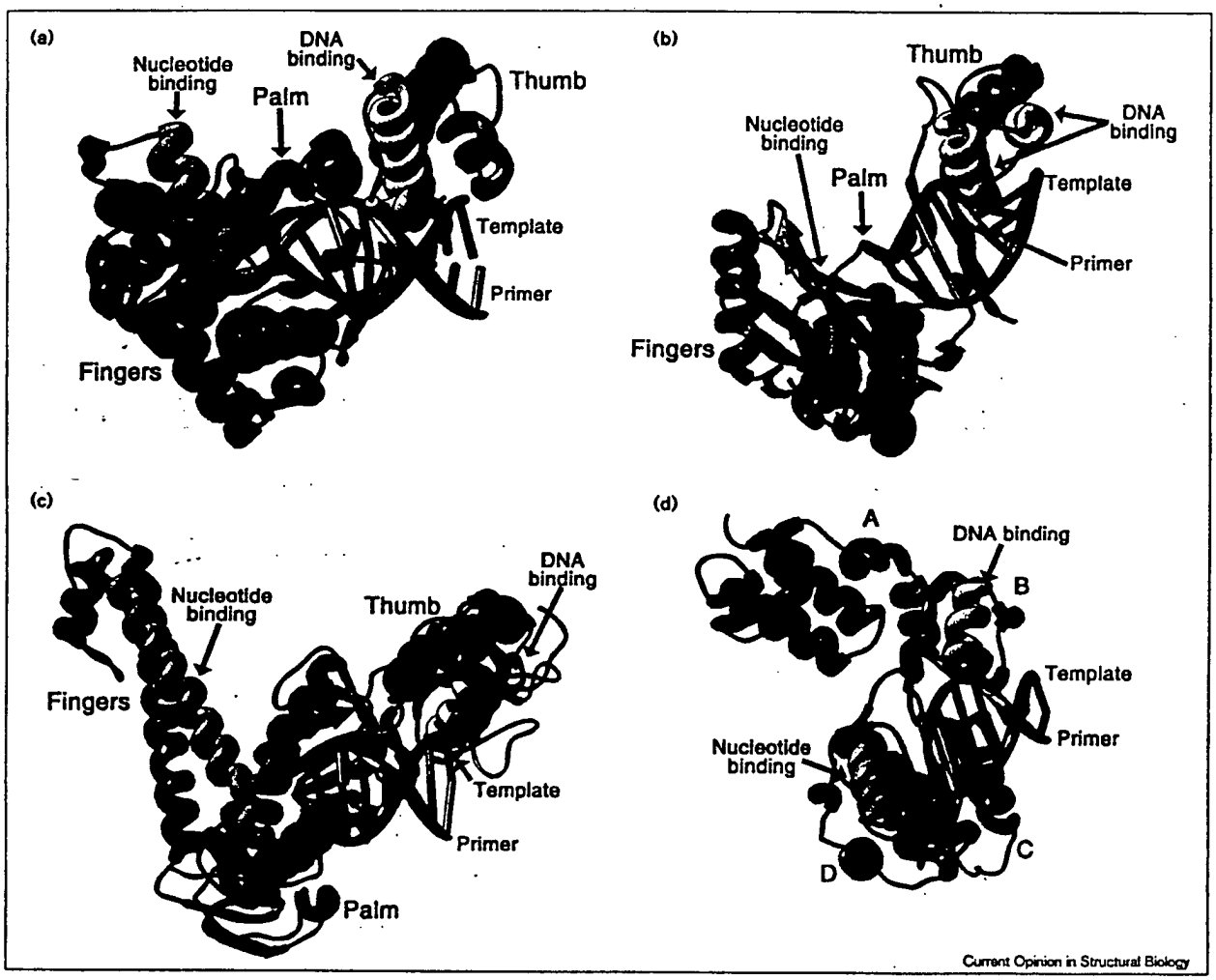
The thumb domain has been implicated in processivity in two other polymerases. When viewed along the bound duplex DNA, KF is seen to surround the DNA on three sides with the tip of the thumb on top [11]. Deleting the tip of the KF thumb reduces its processivity about fourfold [6••]. Furthermore, a thioredoxin-sensitive increase in processivity can be conferred on KF by grafting the thioredoxin-binding loop from T7 pol onto the tip of the KF thumb [37••]. In pol β, deletion of an 8 kDa extension from the tip of its thumb likewise reduces its processivity [38]. Also, a metal ion mediates interactions between the pol β thumb and the substrate DNA, possibly discouraging the DNA from dissociating during the enzyme's catalytic cycle [5•]. Thus, in these three examples (and perhaps in all polymerases), the thumb plays an important role in maintaining the polymerase's grip on its DNA substrates.

Fidelity

An enzyme's fidelity, or its ability to incorporate the correct substrate from among a sea of potential imitators, is a general problem in enzymology. The problem is magnified with DNA polymerases. These enzymes must choose from a pool of four deoxynucleoside triphosphates (dNTPs) and this choice is directed by yet another substrate (the single-stranded template). Furthermore, the consequences of an incorrect choice may be catastrophic to the organism. Thus, fidelity is of central importance in the reaction catalyzed by DNA polymerases. It is therefore surprising that in the T7 DNA polymerase quaternary structure [2••] there are no hydrogen bonds between the protein and the bases of the incipient base pair formed by the template and incoming nucleotide. Instead, it seems that the basis for error discrimination at this juncture is the steric complementarity between the protein and a correctly formed Watson–Crick base pair. This is in contrast to the pol β ternary structure [25], in which polymerase residues hydrogen bond to the nascent base pair. Amino acids from the palm of the T7 pol do, however, interact with base pairs near the 3′ terminus of the primer, suggesting that these residues are scanning for newly-made errors in polymerization. Since mutation of these residues results in a lower affinity of the enzyme for DNA [39,40], they could

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Figure 3



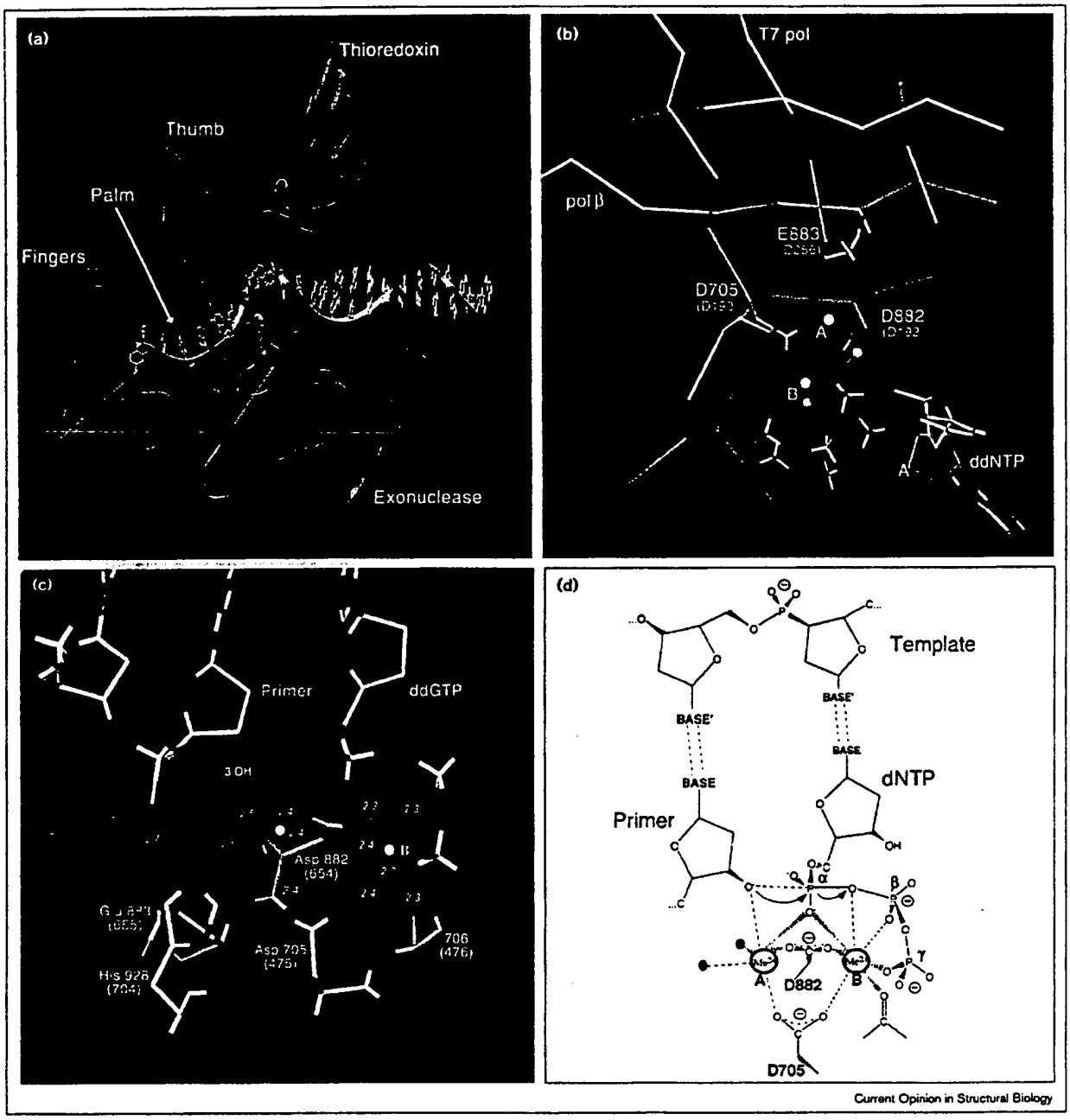
A comparison of primer–template DNA bound to four DNA polymerases. The complexes shown in (a), (b), and (d) are co-crystal structures, whereas the complex in part (c) is a homology model [1••]. These four structures have been aligned by the first two base pairs at the primer terminus. The fingers of these structures are colored blue, the palms are red, and the thumbs are green. Secondary structural elements at the primer terminus that help bind nucleotide are shown in cyan and those that contact DNA are yellow. The primer strand is shown in orange, while the template strand is brown. **(a)** *Taq* pol bound to DNA [4••]. As with the other three structures shown in this figure, the DNA stacks against the fingers and is contacted across the minor groove by the thumb domain. **(b)** The binary complex of HIV-1 RT and DNA [27]. This structure does not have a nucleotide-binding α helix in the fingers domains. Instead, a β hairpin probably performs this function. **(c)** The model of DNA bound to RB69 gp43 [1••]. A likely DNA-binding α helix has been highlighted. It appears that the thumb domain would have to move toward the primer terminus in order to bind DNA analogously to the other polymerases. **(d)** The ternary complex of rat pol β with DNA and ddNTP [25]. Domain D (purple) plays the role of the fingers and presents an α helix at the primer terminus. Domain B is analogous to other polymerase thumb domains and binds the minor groove of the duplex substrate.

be involved in shuttling newly incorporated mismatched nucleotides to the 3′–5′ exonucleolytic site of T7 pol [2••].

Several recently determined structures of human pol β complexed with substrate DNA also address the issue of fidelity. First, polymerases are more error-prone when the native, catalytic Mg^{2+} is replaced by Mn^{2+} [41–43]. Crystalline pol β is capable of adding a nucleotide, untemplated, onto a blunt-ended DNA in the presence of Mn^{2+} (or other mutagenic metal ions), but not in the presence of Mg^{2+} [7•]. This implies that mutagenic cations, perhaps owing to their increased affinity for carboxylates, cause mistakes by stabilizing non-native primer–template–nucleotide combinations. Also, observations of the interactions of pol β with DNA in the crystal have led to a proposal for how this polymerase binds at the single-stranded DNA gap [5•] that is presumably its native substrate. The position of the primer–template in

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Figure 4



Nucleotidyl transfer in the T7 pol co-crystal structure. **(a)** Structure of T7 DNA polymerase complexed with thioredoxin, DNA, and incoming ddNTP. The primer is depicted in magenta and the template is in yellow. White bases denote the portion of the DNA that has been modeled. Purple cylinders represent polymerase α helices and the red cylinders are α helices from thioredoxin. The DNA abuts the fingers at the primer terminus and the thumb contacts DNA across the minor groove. **(b)** Superposition of the active sites of T7 pol and pol β using only the DNAs from the T7 quaternary structure and the pol β ternary structure. For T7 pol, KF numbering (yellow) is used and the pol β residues are labeled in white. The figure shows that the essential carboxylate residues and metal ions are in a similar 3D arrangement, despite the obvious disagreement in the direction of the β strands that make up the palm domains. **(c)** Active site of T7 pol shown using KF numbering, with the T7 pol numbers shown below in parentheses. Metal ions A and B are contacted by two protein residues [D882(654) and D705(475)], the phosphates of the ddNTP and two waters. The putative position of the primer's 3′-OH, which would contact metal ion A, is shown by a red star. Figures provided by Tom Ellenberger. **(d)** Mechanism of nucleotidyl transfer in T7 pol, using the KF numbering system. The active site features two metal ions that stabilize the resulting pentacoordinated transition state. Metal ion A activates the primer's 3′-OH for attack on the α-phosphate of the dNTP. Metal ion B plays the dual role of stabilzing the negative charge that builds up on the leaving oxygen, and chelating the β- and γ-phosphates.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

this model is the same as in the ternary complexes, with the single-stranded template making a sharp turn so that the next double-stranded portion of the DNA

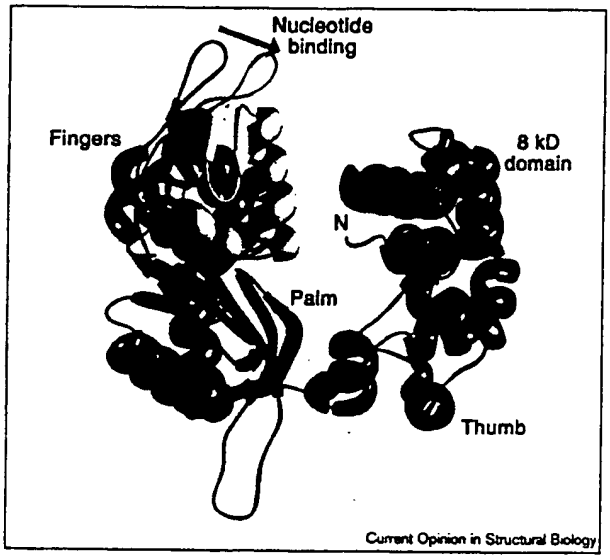
may interact with the 8 kDa domain at the tip of the thumb. Kraut and colleagues [5•] postulate that such a kink in the template has advantages for fidelity, as it would minimize non-specific stacking contacts between the incoming dNTP and the primer; the template strand is also sharply bent in the T7 quaternary complex [2••].

Nucleotidyl transfer

The crystal structures of T7 pol [2••] and of pol β [25] complexed with DNA and next correct dNTP show that these non-homologous enzymes use an identical catalytic mechanism to transfer a nucleotide onto the 3′ end of a DNA primer. The active sites of these unrelated polymerases share many essential and common features. When Ellenberger and colleagues [2••] aligned the two structures by superimposing their DNA substrates, the catalytically essential carboxylates from these disparate structures also superimposed (Figure 4b), as do the two metal ions bound to the triphosphate moiety in both structures. Although the terminal transferase family (to which pol β belongs) has evolved separately from other polymerase families and has a different catalytic domain topology, it shares a common three-dimensional arrangement of critical catalytic components. The structure and positions of the primer–template, dNTP and divalent metal ions are nearly identical in T7 pol and pol β.

The structure of the T7 pol quaternary complex provides a basis for understanding the roles of certain residues known from other studies to be essential for nucleotidyl transfer ([2••]; Figure 4c,4d). For the following discussion, the KF numbering scheme is used. Two conserved carboxylates, D705 and D882, are bridging ligands to the two catalytic metal ions. Significantly, these are the only two carboxylate residues that are absolutely conserved between the pol α, pol I and reverse transcriptase families ([1••]; Figure 2b). E883, which is conserved in the pol I family, does not contact the metal ions, as is consistent with its absence in the pol α family. The two cations have been dubbed A and B. The ligation environment of metal ion A is completed by two water molecules and the *pro*-R oxygen of the nucleotide's α-phosphate. This cation would also be contacted by the 3′-OH of the primer, if it were present (a dideoxy-terminated primer was utilized in the crystallization experiment). This hydroxyl group would be perfectly positioned to undertake an in-line attack on the nucleotide's α-phosphate. Metal ion B is contacted by oxygens from all three of the nucleotide's phosphates, as well as a main-chain oxygen. Curiously, the metal ion–ligand distances in the active site of T7 pol are somewhat longer than expected (Figure 4c). Nevertheless, the features of this Michaelis complex make it clear that a proposed two metal ion mechanism for DNA polymerases [44], which was based on an analogy to the 3′–5′ exonucleolytic two metal ion mechanism of KF [45], is essentially correct (Figure 4d). In this polymerase mechanism, metal ion A activates the 3′-OH of the primer for attack by lowering its pK_a. This nucleophilic attack at the α-phosphate generates a pentacoordinated intermediate, whose structure would be stabilized by both metal ions. A build-up of negative charge on the leaving oxygen of the β-phosphate could be stabilized by metal ion B. Once the reaction is complete, the pyrophosphate product and the metal ions dissociate, and the DNA must translate and rotate relative to the polymerase so that the newly formed primer terminus is correctly positioned in the primer-binding site for another catalytic cycle. The only significant deviation from the earlier proposal is that metal ion B appears to be playing the dual role of transition-state stabilization and chelation of both the β- and γ-phosphates.

Figure 5



Finger domain movement in DNA polymerases. The palm domains of the binary [5•] and ternary [27] complexes of pol β were used to superimpose the two protein structures. The binary (protein and DNA) form of the polymerase is shown in the darker shade and only the fingers domain of the ternary (protein and DNA and dNTP) form is shown in the lighter shade. The fingers close down in the presence of the nucleotide.

Conformational changes

The binding of dNTP and duplex DNA to both pol I family and pol β family DNA polymerases produces large changes in the orientations of the fingers and the thumb domains relative to the palm domain. A comparison of the binary complexes of polymerases and DNA with ternary complexes containing an incoming dNTP shows that the fingers close down on the DNA upon binding of the nucleotide in both pol β ([5•]; Figure 5) and pol-I-like enzymes [2••,4••]. A comparison of the *Taq* and

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

T7 pol complexes shows a few angstroms difference in the position of the primer terminus in these two complexes. Although it is unclear whether these conformational differences are part of the polymerase catalytic cycle or are influenced by crystallographic artifacts, they provide a rational explanation for the kinetic observation [36,46–48] that the rate-limiting step in the polymerase reaction pathway occurs after the binding of dNTP [2••,5•]. The thumb domains of polymerases also move in response to binding DNA [4••,11,25,49], and this mobility may also play a role in the catalytic cycle.

Conclusions

Principal among the many new structural and functional insights into DNA polymerases obtained during the last year have been the structure determinations of a pol α family DNA polymerase and a T7 DNA polymerase complexed with DNA and nucleotide substrates. The structure of RB69 gp43 strongly suggests how the replisomes of this and related polymerases are organized. The structures of *Taq* pol with DNA, pol β with various metal ions and substrates, and the quaternary structure of T7 DNA polymerase all give a much needed insight into the workings of DNA polymerases. These studies, taken together with the results of mutagenic experiments, suggest the functions of the domains that make up polymerases: the fingers are involved in correctly positioning the template and the complementary NTP relative to the catalytic metal ions, and undergo a substrate-induced conformational change that may be important to catalysis; the palm domain harbors two completely conserved carboxylate residues that bind the essential metal ions and also contains residues that may detect replication mistakes; and the thumb domain is important for DNA positioning and processivity. Additionally, the catalytic mechanism of nucleotidyl transfer in the polymerase reaction may be universal, involves two metal ions and is very similar to that of the proof-reading exonucleases. Though the model of RB69 gp43 with DNA and accessory proteins is instructive, it is essential to determine the crystal structures of this and similar polymerases with their substrates and associated factors in order to confirm and extend this model.

Perhaps the most striking findings from these polymerase structures, as anticipated by the limited sequence similarity among families, are the differences in the structures of the fingers and thumb domains of polymerases from different families. The thumb and fingers domain structures are completely unrelated among the four DNA polymerase families now known: pol I, pol α, pol β and reverse transcriptase. Nevertheless they function similarly in all families, using in most cases analogous secondary structural elements. Why are these domains, so essential to the polymerase reaction, derived from different ancestors? It would of interest to establish the structure of the bacterial DNA polymerase III, which shows no recognizable sequence similarity to other polymerases. It may possess yet another polymerase fold and could yield clues to understanding the surprising differences among these vital enzymes.

Acknowledgements

The authors wish to thank S Doublié and T Ellenberger for providing data and figures in advance of publication, J Wang for providing figures and molecular models, and J Jäger and J Pata for access to molecular models. This work was funded in part by the American Cancer Society grant #BE-52K to TAS.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. •• Wang J, Sattar AKMA, Wang CC, Karam JD, Konigsberg WH, Steitz TA: **Crystal structure of a pol α replication DNA polymerase from bacteriophage RB69.** *Cell* 1997, 89:1087-1089.

The authors describe the first crystal structure of either a pol α or 'B family' polymerase. This structure has wide-ranging implications for the relatedness of DNA polymerases, proof-reading, the organization of the RB69 replisome and eukaryotic DNA polymerases that could not be modeled on the pol I enzyme.

2. •• Doublié S, Tabor S, Long AM, Richardson CC, Ellenberger T: **Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution.** *Nature* 1998, 391:251-258.

The elucidation of the structure of the ternary complex of T7 DNA polymerase with its substrates is described. This 2.2 Å resolution structure shows how substrates bind to this KF-like polymerase and suggests a mechanism by which thioredoxin allows the DNA to stay associated with the polymerase. A detailed view of the active site shows that the catalytic mechanism involves two metal ions.

3. • Kiefer JR, Mao C, Hansen CJ, Basehore SL, Hogrefe HH, Braman JC, Beese LS: **Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 Å resolution.** *Structure* 1997, 5:95-108.

The high-resolution structure of the large fragment of this enzyme shows that the polymerase domain of pol I from *B. stearothermophilus* is shown to be highly homologous to the same domain of KF. The structural basis for the lack of 3′–5′ exonuclease activity of the *Bacillus* enzyme is also discussed.

4. •• Eom SH, Wang J, Steitz TA: **Structure of *Taq* polymerase with DNA at the polymerase active site.** *Nature* 1998, 382:278-281.

The structure of a binary complex of *Taq* pol and blunt-ended DNA shows that the DNA binds with the blunt end stacked against the fingers and the double-stranded portion leaving the active site between the thumb and the vestigial exonuclease domain.

5. • Pelletier H, Sawaya MR, Wolfle W, Wilson S, Kraut J: **Crystal structures of human DNA polymerase β complexed with DNA: implications for catalytic mechanism, processivity, and fidelity.** *Biochemistry* 1996, 35:12742-12761.

The authors describe several binary complexes of pol β and DNA. These complexes reveal facts about the processivity, fidelity and the conformational changes upon nucleotide binding. In addition, a model for the binding of pol β to gapped DNA is proposed and discussed in these contexts.

6. •• Minnick DT, Astatke M, Joyce CM, Kunkel TA: **A thumb subdomain mutant of the large fragment of *Escherichia coli* DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity.** *J Biol Chem* 1996, 271:24954-24961.

The results in this paper show that deleting 24 residues from the tip of the thumb of KF reduce the processivity of this enzyme. Further, such a polymerase mutant makes far more frameshift mutations than wild-type KF.

7. • Pelletier H, Sawaya MR, Wolfle W, Wilson SH, Kraut J: **A structural basis for metal ion mutagenicity and nucleotide selectivity in human DNA polymerase β.** *Biochemistry* 1996, 35:12762-12777.

A variety of structures are presented that illustrate the importance of the identity of the metal ions and the base, ribose and triphosphate moieties on the binding and utilization of nucleotides by human pol β.

8. Komberg A, Baker TA: *DNA Replication*, edn 2. New York: WH Freeman; 1992.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

9. Ito J, Braithwaite DK: Compilation and alignment of DNA polymerase sequences. *Nucleic Acids Res* 1991, 19:4045-4057.

10. Wang J, Yu P, Lin TC, Konigsberg WH, Steitz TA: Crystal structures of an NH_2- terminal fragment of T4 DNA polymerase and its complexes with single-stranded DNA and with divalent metal ions. *Biochemistry* 1996, 35:8110-8119.

11. Beese LS, Derbyshire V, Steitz TA: Structure of DNA polymerase I Klenow fragment bound to duplex DNA. *Science* 1993, 260:352-355.

12. Huang WM, Lehman IR: On the exonuclease activity of phage T4 deoxyribonucleic acid polymerase. *J Biol Chem* 1972, 247:3139-3146.

13. Capson TL, Peliska JA, Kaboord BF, Frey MW, Lively C, Dahlberg M, Benkovic SJ: Kinetic characterization of the polymerase and exonuclease activities of the gene 43 protein of bacteriophage T4. *Biochemistry* 1992, 31:10984-10994.

14. Lin TC, Karam G, Konigsberg WH: Isolation, characterization, and kinetic properties of truncated forms of T4 DNA polymerase that exhibit 3′–5′ exonuclease activity. *J Biol Chem* 1994, 269:19286-19294.

15. Sattar AKMA, Lin TC, Jones C, Konigsberg WH: Functional consequences and exonuclease kinetic parameters of point mutations in bacteriophage T4 DNA polymerase. *Biochemistry* 1996, 35:16621-16629.

16. Kuchta RD, Benkovic P, Benkovic SJ: Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. *Biochemistry* 1988, 27:6716-6725.

17. Kong X-P, Onrust R, ODonnell M, Kuriyan J: Three-dimensional structure of the β- subunit of *E. coli* DNA polymerase III holoenzyme: a sliding clamp. *Cell* 1992, 69:425-437.

18. Berdis LS, Soumillion P, Benkovic SJ: The carboxyl terminus of the bacteriophage T4 DNA polymerase is required for holoenzyme complex formation. *Proc Natl Acad Sci USA* 1996, 93:12822-12827.

19. Delarue M, Poch O, Tordo N, Moras D, Argos P: An attempt to unify the structure of polymerases. *Protein Eng* 1990, 3:461-467.

20. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA: Crystal structure at 3.5 Å of HIV-1 reverse transcriptase complexed with an inhibitor. *Science* 1992, 256:1781-1790.

21. Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA: Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. *Nature* 1985, 313:762-766.

22. Kim Y, Eom SH, Wang J, Lee DS, Suh SW, Steitz TA: Crystal structure of *Thermus aquaticus* DNA polymerase. *Nature* 1995, 376:612-616.

23. Korolev S, Nayal M, Barnes WM, DiCera E, Waksman G: Crystal structure of the large fragment of *Thermus aquaticus* DNA polymerase I at 2.5 Å resolution: structural basis for thermostability. *Proc Natl Acad Sci USA* 1995, 92:9264-9268.

24. Davies JF Jr, Almassy RJ, Hostomska Z, Ferre RA, Hostomsky Z: 2.3 Å crystal structure of the catalytic domain of DNA polymerase β. *Cell* 1994, 76:1123-1133.

25. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J: Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. *Science* 1994, 264:1891-1903.

26. Holm L, Sander C: DNA polymerase β belongs to an ancient nucleotidyl transferase superfamily. *Trends Biochem Sci* 1995, 20:345-347.

27. Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Jr, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Clark P *et al.*: Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA. *Proc Natl Acad Sci USA* 1993, 90:6320-6324.

28. Ren J, Esnouf R, Hopkins A, Ross C, Jones Y, Stammers D, Stuart D: The structure of HIV-1 reverse transcriptase complexed with 9-chloro-TIBO: lessons for inhibitor design. *Structure* 1995, 3:915-926.

29. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison, SC. The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1. *Proc Natl Acad Sci USA* 1995, 92:1222-1226.

30. Georgiadis MM, Jessen SM, Ogata CM, Telesnitsky A, Goff SP, Hendrickson WA: Mechanistic implications from the structure of a catalytic fragment of Moloney murine leukemia virus reverse transcriptase. *Structure* 1995, 3:879-892.

31. Steitz, TA, Smerdon SJ, Jäger J, Joyce CM: A unified mechanism for nonhomologous DNA and RNA polymerases. *Science* 1994, 266:2022-2025.

32. Pelletier H: Polymerase structures and mechanism. *Science* 1994, 266:2025-2026.

33. Hughes SH, Hostomsky Z, LeGrice SFJ, Lentz K, Arnold E: What is the orientation of DNA polymerases on their templates? *J Virol* 1996, 70:2679-2683.

34. Modrich P, Richardson CC: Bacteriophage T7 deoxyribonucleic acid replication *in vitro*. A protein of *Escherichia coli* required for bacteriophage T7 DNA polymerase activity. *J Biol Chem* 1975, 250:5508-5514.

35. Kati WM, Johnson KA, Jerva LF, Anderson KS: Mechanism and fidelity of HIV reverse transcriptase. *J Biol Chem* 1992, 267:25988-25997.

36. Werneberg BG, Ahn J, Zhong X, Hondal RJ, Kraynov VS, Tsai M-D: DNA polymerase β: pre-steady-state kinetic analysis and roles of arginine-283 in catalysis and fidelity. *Biochemistry* 1996, 35:7041-7050.

37. •• Bedford E, Tabor S, Richardson CC: The thioredoxin binding domain of bacteriophage T7 DNA polymerase confers processivity on *Escherichia coli* DNA polymerase I. *Proc Natl Acad Sci USA* 1997, 94:479-484.

By simply grafting the thioredoxin-binding domain of T7 DNA polymerase onto the tip of the thumb of a KF, the authors show that the KF can be rendered into a thioredoxin-sensitive processive DNA polymerase. Though the overall activity of the mutant KF is lower than the wild-type, the gains in processivity are impressive.

38. Singhal RK, Wilson SH: Short gap-filling synthesis by DNA polymerase β is processive. *J Biol Chem* 1993, 268:15906-15911.

39. Polesky AH, Steitz TA, Grindley NDF, Joyce CM: Identification of residues critical for the polymerase activity of the Klenow fragment of DNA polymerase I from *Escherichia coli*. *J Biol Chem* 1990, 265:14579-14591.

40. Polesky AH, Dahlberg ME, Benkovic SJ, Grindley NDF, Joyce CM: Side chains involved in catalysis of the polymerase reaction of DNA polymerase I from *Escherichia coli*. *J Biol Chem* 1992, 267:8417-8428.

41. El-Deiry WS, Downey KM, So AG: Molecular mechanisms of manganese mutagenesis. *Proc Natl Acad Sci USA* 1984, 81:7378-7382.

42. Orgel A, Orgel LE: Induction of mutations in bacteriophage T4 with divalent manganese. *J Mol Biol* 1965, 14:453-457.

43. Beckman RA, Mildvan AS, Loeb LA: On the fidelity of DNA replication: manganese mutagenesis *in vitro*. *Biochemistry* 1985, 24:5810-5817.

44. Steitz TA: DNA- and RNA-dependent DNA polymerases. *Curr Opin Struct Biol* 1993, 3:31-38.

45. Beese LS, Steitz TA: Structural basis for the 3′–5′ exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. *EMBO J* 1991, 10:25-33.

46. Bryant FR, Johnson KA, Benkovic SJ: Elementary steps in the DNA polymerase I reaction pathway. *Biochemistry* 1983, 22:3537-3546.

47. Patel SS, Wong I, Johnson KA: Pre-steady-state kinetic analysis of processive DNA replication including complete characterization of an exonuclease-deficient mutant. *Biochemistry* 1991, 30:511-525.

48. Spence RA, Kati WM, Anderson KS, Johnson KA: Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors. *Science* 1995, 267:988-993.

49. Jäger J, Smerdon SJ, Wang J, Boisvert DC, Steitz TA: Comparison of three different crystal forms shows HIV-1 reverse transcriptase displays an internal swivel motion. *Structure* 1994, 2:869-876.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

...lar Structures
...on available now
See inside front cover for details

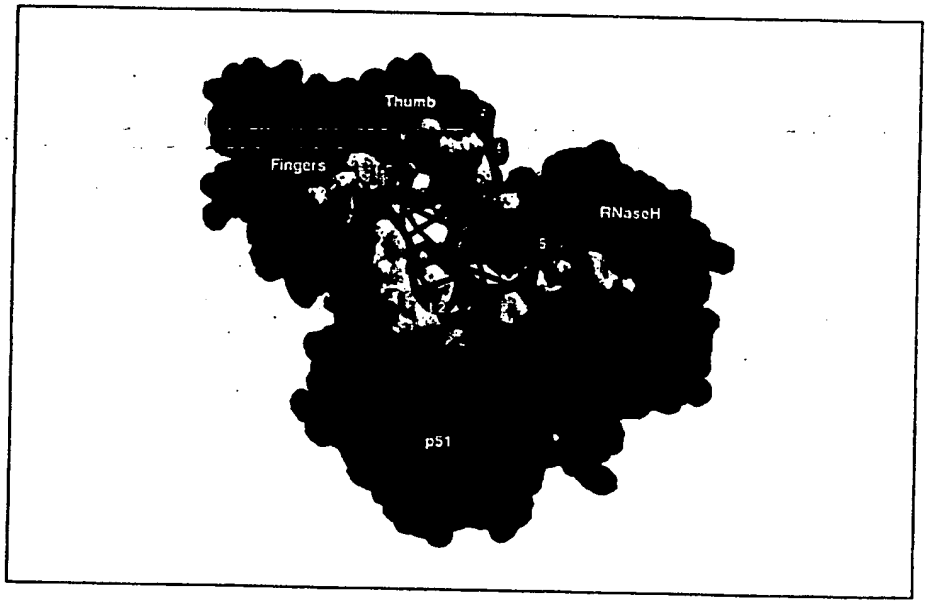
Structural Biology

Protein–nucleic acid interactions
Edited by Simon EV Phillips and Dino Moras

Folding and binding
Edited by Christopher M Dobson and Oleg B Ptitsyn

Paper alert

Web alert



This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Current Opinion in **Structural Biology**

Current Opinion in Structural Biology
ISSN 0959-440X
is published bimonthly by

Current Biology Publications
34–42 Cleveland Street
London W1P 6LE, UK

Each volume consists of six issues of approximately 130 pages.

London office
34–42 Cleveland Street
London W1P 6LE, UK

In-house Editor Rachel Page
e-mail: rachelp@current-biology.com
Editorial Administrator Samantha Bedser
Editorial Assistant Clare Rathbone
Managing Editor Julie Solomon
Editorial Director David Weedon
Illustrator Jamie Oliver
Print production Yolanda Perez Capilla
Electronic production Fehim Osman
Database Patrick Morgan
Marketing Nik Screen

Advertising Deborah Breeds
Tel: +44 (0)171 323 5345
Fax: +44 (0)171 436 9293
e-mail: adscb@current-biology.com

Copyright © 1999 by Elsevier Science Ltd
No part of this publication may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic or otherwise, without prior permission of the copyright owner. Authorization to photocopy items from this publication for personal or internal use, or for the personal or internal use of specific clients, is granted by Current Biology Publications on the condition that the copier pay the Copyright Clearance Center Inc. (CCC) the base fee of $1.00 per journal page per copy. This consent does not extend to multiple copying for promotional purposes. Payment should include the fee code ISSN 0959-440X/99, $1.00 per journal page per copy and should be made directly to CCC, 222 Rosewood Drive, Danvers, MA 01293, USA. For all other use, permission should be sought directly from

Petra Maraj
Current Biology Publications
34–42 Cleveland Street
London W1P 6LE, UK

Tel: +44 (0)171 580 8377
Fax: +44 (0)171 580 8428
e-mail: petra@current-biology.com

The Aims of the journal can be found at http://biomednet.com/cbiology/journals.htm

Printed on acid-free paper by Bell and Bain Ltd, Glasgow, UK.

Current Opinion in Structural Biology is indexed and/or abstracted by Biosis, Elsevier BIOBASE / Current Awareness in Biological Sciences, Current Contents / Life Sciences, EMBASE/Excerpta Medica, Index Medicus / MEDLINE, Reference Update.

Whilst every effort is made by the publishers and editorial board to see that no inaccurate or misleading data, opinion or statement appear in this journal, they wish to make it clear that the data and opinions appearing in the articles and advertisements herein are the responsibility of the contributor or advertiser concerned. Accordingly, the publishers, the editorial board and section editors and their respective employees, officers and agents accept no liability whatsoever for the consequences of any such inaccurate or misleading data, opinion or statement.

Subscription rates 1999, six issues. Print prices include airspeed delivery.

	Personal		Institutional		Student[3]	
	North America[2]	Rest of World	North America[2]	Rest of World	North America[2]	Rest of World
Print only	$206	£125	$921	£560	$99	£60
Online only[1]	$227	£138	Not available	Not available	Not available	Not available
Print + online[1]	$247	£150	Not available	Not available	$99	£60

[1]EU subscribers will be liable to VAT on online subscriptions and may be liable to VAT on print subscriptions.
[2]Canadian subscribers must add GST.
[3]Students must give the name of their institution or school plus the name of their department chairman to qualify for the student rate, which is available for a maximum of two years.

Subscription orders and enquiries
should be sent directly to
Current Biology Publications
PO Box 331
Haywards Heath
West Sussex RH16 3FG, UK

Tel: +44 (0)1444 475650
Fax: +44 (0)1444 445423
e-mail: cb.subs@rbi.co.uk

Or contact Elsevier Science's Regional Sales Offices:

From USA and Canada, contact the New York office
Tel: (+1) 212 633 3730
Toll Free number for North American customers:
1-888-4ES-INFO (437-4636)
Fax: (+1) 212 633 3680
e-mail: usinfo-f@elsevier.com

From Europe, Middle East and Africa, contact the Amsterdam office
Tel: (+31) 20 485 3757
Fax: (+31) 20 485 3432
e-mail: nlinfo-f@elsevier.nl

From Asia, contact the Singapore office
Tel: (+65) 434 3727
Fax: (+65) 337 2230
e-mail: asiainfo@elsevier.com.sg

From Japan, contact the Tokyo office
Tel: (+81) 3 5561 5033
Fax: (+81) 3 5561 5047
e-mail: info@elsevier.co.jp

From Latin America, contact the Rio de Janeiro office
Tel: (+55) 21 509 5340
Fax: (+55) 21 507 1991
e-mail: elsevier@campus.com.br

From Australasia and NZ, contact Sheila Mayer
Tel: (+61) 029 958 4429
Fax: (+61) 029 967 2126
e-mail: elsevier@ozemail.com.au

Send notices of change of address at least 8 weeks in advance, including both old and new addresses. Cancellations on renewed subscriptions will not be accepted after the first issue has been shipped.

Online subscriptions can be obtained through the BioMedNet library at http://biomednet.com

In case of difficulty please contact
Nik Screen
Current Biology Publications
34–42 Cleveland Street
London W1P 6LE, UK

Tel: +44 (0)171 580 8377
Fax: +44 (0)171 580 8428
e-mail: marketing@current-biology.com

Periodicals Class postage paid at Middlesex NJ.
US postal service registration number 006-611

US Postmaster:
Airfreight and mailing in the USA
c/o Pronto Mailers
P.O. Box 177
Middlesex
NJ 08846, USA

Send changes of address to
Current Opinion in Structural Biology
c/o Elsevier Science
Regional Sales Office
Customer Support Department
PO Box 945
New York, NY 10159-0945, USA

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Getting a grip: polymerases and their substrate complexes

Joachim Jäger* and Janice D Pata†

Underpinned by a database of more than a dozen different crystal structures, an increasingly complete and coherent picture of polymerase structure and function is emerging. Recently determined structures of DNA and RNA polymerases have revealed some of the molecular features and structural changes governing catalysis, oligomerization, processivity and fidelity. Despite having minimal similarities in sequence and protein topology, the polymerases all display a functionally analogous set of subdomains that bind the primer, template and nucleotide substrates in similar though not identical fashions. The two-metal-ion mechanism for nucleotide incorporation, however, is shared even by nonhomologous polymerases.

Addresses
*School of Biochemistry and Molecular Biology, University of Leeds, Leeds, LS2 9JT, UK; e-mail: jj@bmb.leeds.ac.uk
†Department of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT 06520, USA

Current Opinion in Structural Biology 1999, 9:21–28

http://biomednet.com/elecref/0959440X00900021

© Elsevier Science Ltd ISSN 0959-440X

Abbreviations

3Dpol	poliovirus 3D RNA polymerase
BF	*B. stearothermophilus* fragment
DNAP	DNA polymerase
ds	double-stranded
KF	Klenow fragment
NNI	non-nucleoside inhibitor
PDB	Protein Data Bank
pol β	DNA polymerase β
pol I	polymerase I
RNAP	RNA polymerase
RT	reverse transcriptase

Introduction

Replication and transcription are of vital importance to all living organisms. These processes, catalyzed by the polynucleotide polymerases, must be carried out in a precise and efficient manner. The past few years have seen an explosive growth in the number of polymerase structures determined and a corresponding growth in the understanding of polymerase activities. This review will focus primarily on a selected subset of the most recent structures determined from each of the four major polymerase classes (Table 1). These structures include three DNA-dependent DNA polymerases [1•–3•] in complexes with substrates and products, an RNA-dependent DNA polymerase [4•] and a DNA-dependent RNA polymerase [5•] bound, respectively, to RNA and protein inhibitors, and also the apoenzyme structure of an RNA-dependent RNA polymerase [6•].

The crystal structure of the Klenow fragment (KF) of *Escherichia coli* polymerase I (pol I) was the first polymerase structure determined [7]. The overall shape of the polymerase domain was likened to a right hand, with subdomains termed fingers, palm and thumb. This analogy has proven useful for all polymerase structures solved to date. A comparison with the structure of kanamycin nucleotidyl transferase [8,9] shows distinct differences in the palm subdomains, indicating that the structures belong to two evolutionarily distinct groups [10]. Although we will refer to these groups as the polymerase family and the nucleotidyl transferase family, both of these evolutionary families belong functionally to the nucleotidyl transferase enzyme class (Enzyme Commission number 2.7.7.7). With the exception of DNA polymerase β (pol β), all the polymerase structures determined to date belong to the polymerase family. Nucleotidyl transferases are not responsible for genome replication and transcription, rather they perform more specialized tasks, such as DNA repair and RNA maturation.

DNA-dependent DNA polymerases

The most thoroughly studied polymerases are those of the pol I family [11–13]; also referred to as 'family A' DNA polymerases (DNAPs) [14]. Since the structure of the KF was determined in 1985 [7], the structures of three homologous enzymes have been solved. These structures, from bacteriophage T7 (T7 DNAP [1•]), *Thermus aquaticus* (Taq DNAP [15,16]) and *Bacillus stearothermophilus* (BF [2•,17]), show that the polymerase domains of the pol I enzymes are nearly identical to each other (Figure 1). In addition to the pol I structures, the important structure of gp43 from bacteriophage RB69, a member of the polymerase α family (also referred to as the family B polymerases [14]) of replicative polymerases, has been recently determined [18], but will not be discussed here as it has been thoroughly reviewed previously [19].

Although a structure of KF with DNA duplex located at its polymerase active site remains elusive, several such complexes of other pol I enzymes have been obtained (Table 1). In particular, the high resolution structures of T7 DNAP and BF complexed with nucleic acids give more insights into the structure and mobility of the polymerase subdomain, as well as unprecedented structural details about substrate and cofactor binding in the pol I family. In the 2.2 Å structure of T7 DNAP, the enzyme, bound to its processivity factor *E. coli* thioredoxin, is trapped at a stage just prior to phosphoryl transfer. The complex contains dideoxy-terminated primer–template DNA, the next incoming nucleotide and two magnesium ions. The 1.8 Å BF structures show primer–template duplexes before and after the incorporation of nucleotides within the crystal.

The catalytic core of the polymerase palm subdomain is a common folding unit and it appears in many different structural contexts [6•,20,21]. It consists of a three-stranded antiparallel β sheet flanked by two α helices. In

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Table 1

Recently determined polymerase structures.

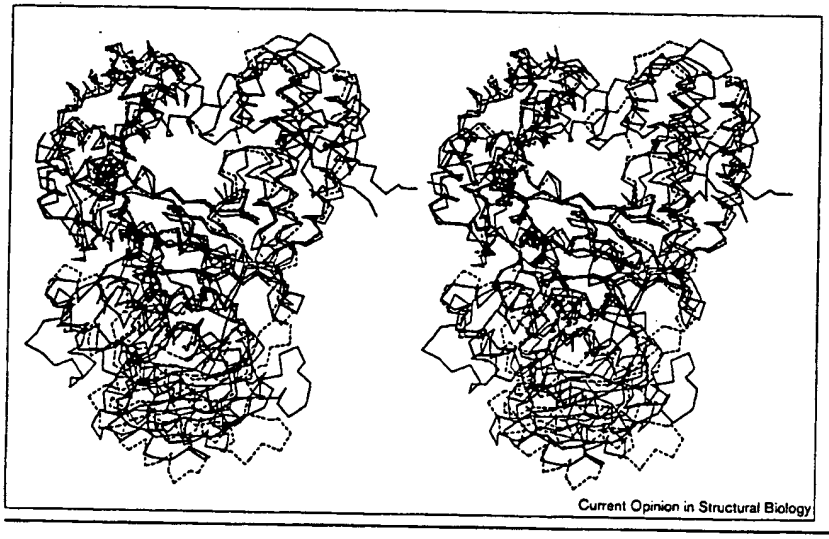
Organism	Class	Template	Size	Complex formed	Resolution (Å)	PDB code	Reference
E. coli	DNA pol	DNA	Fragment	*p*-thio-thymidine and zinc	2.1	1krp,1ksp	[24]
T. aquaticus	DNA pol	DNA	Full	dsDNA (8/8-mer)	2.8	1tau	[16]
T. aquaticus	DNA pol	DNA	Fragment	dCTP	2.5	5ktq	[57]
B. stearothermophilus (II)	DNA pol	DNA	Full	Apo	2.1	1bdp	[17]
B. stearothermophilus (II)	DNA pol	DNA	Fragment	dsDNA (11/9-mer) and magnesium	1.8	2bdp	[2•]
RB69 phage	DNA pol	DNA	Full	GMP	2.8	1waj	[18]
T7 phage	DNA pol	DNA	Full	dsDNA (13/10-mer), magnesium, ddGTP and *E. coli* thioredoxin	2.2	1t7p	[1•]
Human	DNA NT	DNA	Full	Gapped dsDNA(16/9/5-mer), ddCTP and magnesium	2.2	1bpy	[3•]
HIV-1	DNA pol	RNA or DNA	Full	RNA pseudoknot (33-mer)	4.8	1huv	[4•]
T7 phage	RNA pol	DNA	Full	T7 lysozyme	2.8	1aro	[5•]
E. coli	RNA pol	DNA	Fragment	Apo	2.5	1bdf	[56]
Poliovirus	RNA pol	RNA	Partial	Apo (calcium ions)	2.6	1rdr	[6•]

NT, nucleotidyl transferase.

polymerases, this region contains the most highly conserved residues and is used as a binding platform for the DNA primer, dNTPs and the critical divalent metal ions, which are bound by a network of charged interactions. In the ternary complex of T7 DNAP, for example, metals A and B are coordinated by two strictly conserved aspartate sidechains in sequence motifs A and C [12,18], a carbonyl oxygen, two water molecules and by the α, β and γ phosphate oxygens of the incoming NTP. A hydroxyl group at the 3′ end of the primer would be located within 3.1 Å of the α phosphorous atom and thus would be capable of an in-line attack (see Figure 3b in [1•]). Even though a water molecule appears to occupy the position of metal A in BF polymerase, the active site geometry remains very similar to that observed in T7 DNAP. As had been predicted earlier [22,23], the two ions thus serve multiple purposes: compensation for the negative charges both on strictly conserved acidic residues and on the incoming NTP; stabilization of the negative charge on the attacking 3′ hydroxyl oxygen atom; and stabilization of the pentacovalent transition state [1•–3•,19,24].

In all recently determined pol I structures, the thumb and fingers subdomains are found to be in close contact ([15,17,25]; J Jäger, unpublished data). Thus, the apparently large polymerase cleft that had been proposed to be the pathway for the growing primer–template [7] is too narrow to accommodate duplex DNA. The narrow channel that remains between the fingers and thumb presumably enables the entry of NTPs and the exit of pyrophosphate

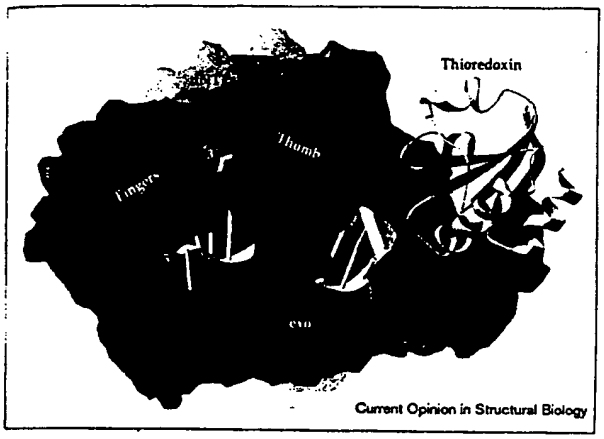
Figure 1



Stereo superposition of DNA polymerases from *E. coli* (dashed blue lines), *B. stearothermophilus* (red lines) and T7 bacteriophage (green lines). The overall structures of the polymerase domains are strikingly well conserved among the different members of the pol I subfamily. The sequence identities based upon the structural alignments shown are only 46% and 31% for BF and T7 DNAP, respectively. The α carbons of the KF polymerase domain (residues 545–928) superimpose on equivalent atoms in T7 DNAP and BF with root mean square deviations of 1.96 Å (261 matching residues) and 1.26 Å (333 matching residues), respectively. Note the large differences in the exonuclease domains (bottom half).

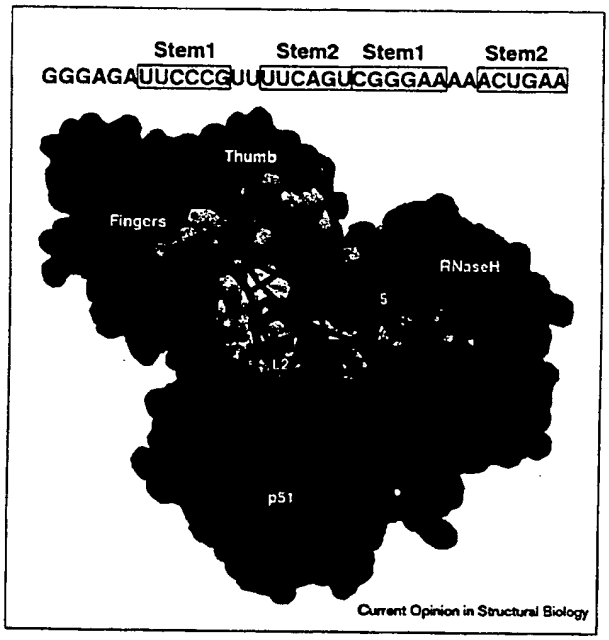
This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Figure 2



Structure of T7 DNA polymerase complexed with dsDNA and *E. coli* thioredoxin. The polymerase domain is viewed from above into the polymerase active site, showing the 3′ end of the primer strand. The fingers and thumb domains are in close contact. The two protrusions on the thumb domain, together with the large exonuclease domain (exo), almost surround the DNA duplex. The template strand does not enter the active site through the cleft between the fingers and thumb domains; instead it traverses the fingers domain. The orientation of the primer–template is very similar in all pol I like enzymes. Furthermore, it seems plausible that this orientation is preserved in all other polymerases as a result of the conservation of the core of the palm domains.

Figure 3



Interaction of an RNA pseudoknot inhibitor (sequence given at top) with HIV-1 RT. The molecular surface rendering of HIV-1 RT indicates the contact regions with the RNA. Surface residues less than 5 Å away from the inhibitor are shown in the lightest gray. Stem S2 at the 3′ end of the pseudoknot forms strong electrostatic interactions with residues in the thumb and fingers domains. The 5′ end is found in the vicinity of the RNaseH active site. Loop L2 (foreground) interacts with the p51 connection domain.

after catalysis (Figure 2). In the KF, this channel is lined with conserved residues His734, Arg754, Lys758, Tyr766 and, on the opposite side, conserved residues Lys635 and Arg682. The X-ray structures of Taq DNAP, T7 DNAP and BF complexed with DNA reveal that the 3′ end of the primer is positioned at the polymerase active site, with the duplex clamped down by the thumb and the 5′ end of the template stand continuing past the active site and traversing the fingers. The clamping of the duplex is enhanced by the tip of the thumb. This structural feature is found in all pol I-like polymerases and may well control the processivity of these enzymes, as they usually work without auxiliary processivity factors [26]. In T7 DNAP, the thumb extends more than 20 Å over the nucleic acid backbone, thereby surrounding more than 60% of the duplex (Figure 2).

The junction between the palm and fingers subdomains plays a crucial role in the precise formation and recognition of canonical Watson–Crick base pairs. Together with the terminal base pair of the primer–template, residues from helices O and Q in the pol I polymerases form a tightly constrained binding pocket that is unsuited for mismatched base pairs [1•]. Distortions are sensed by a range of van der Waals', stacking and hydrogen bond interactions. The residues involved in mismatch recognition are strictly conserved within the pol I subfamily. In KF, these residues are Arg668, Tyr766, Gly767, Asn845, Gln849 and Lys857. Residues corresponding to those from helix Q in KF have been identified even in the RNA-dependent RNA polymerases, following considerable structure-based realignment [6•,27,28]. This conservation may indicate that the elaborate mechanisms governing template recognition and fidelity in pol I-like polymerases may function in other polymerase classes as well.

In T7 DNAP, the formation of the recognition pocket is accompanied by a previously unsuspected tightening of the protein matrix around the terminal base pair, through the rotation of helices O, O1, O2 and P (fingers) towards the substrates. Similar motions may occur during the catalytic cycle of Taq DNAP and BF, but are not observed in the binary Taq DNAP–DNA and BF–DNA co-crystal structures, presumably because they are not trapped at the stage of nucleotide incorporation. These complexes resemble preincorporation or post-translocation complexes. In contrast to the T7 DNAP and polymerase β ([3•], discussed below) co-crystal structures, the 3′ end of the primer in the BF complex is not terminated and, thus, is available for turn-over. Difference Fourier maps of BF complexes at near atomic resolution, before and after the addition of dNTPs, clearly show that the enzyme incorporated the next correct nucleotide. Ultimately, time-resolved X-ray crystallography may allow us to follow catalysis within the crystalline environment, provided that lattice interactions do not interfere with the motions of the fingers subdomain or the primer–template.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Nucleotidyl transferases

DNA polymerase β (pol β) participates in the mammalian base excision DNA repair pathway. Crystal structures of rat DNA polymerase β showed that it has two domains — an N-terminal 8 kDa domain with DNA lyase activity and a C-terminal 31 kDa domain with nucleotidyl transferase activity [20,29,30]. Like the pol I polymerase domain, the nucleotidyl transferase domain was described using analogy to a right hand. For the analogy to be useful, however, we have swapped the fingers and thumb designations, so that analogous structures serve analogous functions [31]. Thus, the thumb subdomain of pol β is adjacent to the 8 kDa domain and the fingers subdomain is at the C terminus of the protein.

The *in vivo* activity of pol β has been elucidated in three recently determined structures [3•]. These structures are of human pol β complexed with either a one-nucleotide gapped DNA duplex, a dideoxynucleotide-terminated gapped duplex and incoming ddCTP, or nicked DNA that would be the result of a single nucleotide incorporation into gapped DNA. The structures show that when the natural gapped DNA substrate is bound, the 8 kDa domain reaches from the thumb subdomain over the active site cleft, which is perhaps analogous to the contact between the fingers and thumb of pol I enzymes. The DNA template strand has a 90° kink located to the 3′ side of the template base and the downstream DNA duplex traverses the fingers and interacts with the 8 kDa domain. When the incoming nucleotide is in place, the fingers and 8 kDa domains close down more tightly around the DNA, forming the active conformation of the catalytic site. Two magnesium ions are found in the active site. One ion, metal B, forms tridentate contacts to the α, β and γ phosphates of the incoming dNTP. With respect to the primer–template and dNTP, metal A is positioned as in the T7 DNAP structure, in order to activate the 3′ hydroxyl of the primer strand. The base of the incoming dNTP fits into a pocket formed by residues of the fingers and the terminal base pair and is positioned such that the α phosphate is aligned with the 3′ hydroxyl of the primer strand. The protein also forms hydrogen bonds and van der Waals' interactions with the minor groove of the newly formed base pair. Both the minor groove interactions and nucleotide-binding pocket select for a correctly formed Watson–Crick base pair.

RNA-dependent DNA polymerases

The first structure of an RNA-dependent DNA polymerase, that of HIV-1 reverse transcriptase (RT) complexed with the non-nucleoside inhibitor (NNI) Nevirapine, was described in 1992 [32] and was subsequently refined to 2.9 Å resolution [33]. The three-dimensional structure reveals a strikingly asymmetric heterodimer consisting of two differently folded subunits (molecular weights 66 kDa and 51 kDa). The p66 subunit is characterized by a large polymerase cleft composed of subdomains named, in analogy to the pol I enzyme family, fingers, palm and thumb. A connection domain links the polymerase domain to a RNaseH domain. The p51 subunit has the same sequence as the N-terminal 440 residues of p66 [34], but it lacks the RNaseH domain. High resolution crystal structures of RT fragments have also been reported by Unge *et al.* [35] (residues 1–216 of HIV-1 RT, Protein Data Bank [PDB] code 1har) and Georgiadis *et al.* [21] (residues 10–278 of Moloney murine leukemia virus RT, PDB code 1mml). A comparison of several RT crystal structures shows that the protein appears to have a specific flexibility that allows the rotation of the polymerase domain relative to the rest of the molecule [36]. The observed swiveling motion may allow RT to accommodate the rotational and translational movements of the growing nucleic acid duplex, which present an especial problem for RT because it uses an asymmetric molecule (tRNA Lys3) as the primer for first strand synthesis.

The crystal form of HIV-1 RT first obtained by Kohlstaedt *et al.* [32] served to elucidate at medium resolution the co-crystal structures of other NNI complexes [37]. The report of high resolution structures (to 2.2 Å) by Stuart and co-workers [38,39] was a major breakthrough, however, as it provided a very detailed description of the NNI-binding pocket. Until then, inaccuracies and the lack of solvent molecules near the NNI-binding site in the previously reported models had hampered structure-based drug design studies.

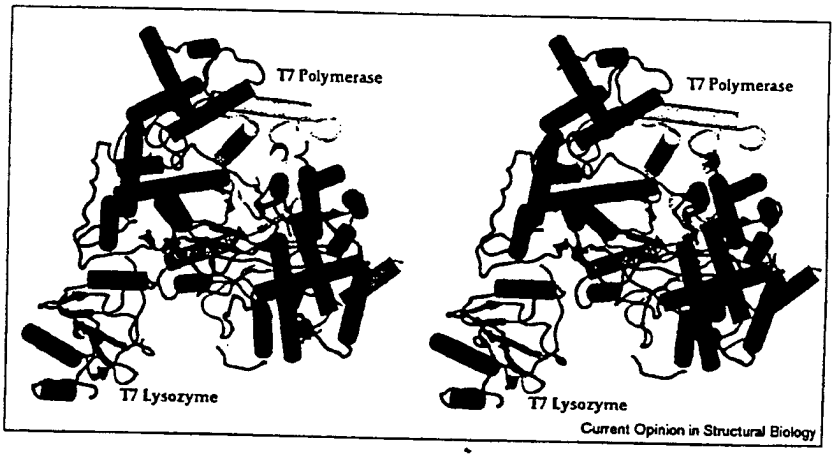
A complex of HIV-1 RT with a double-stranded (ds)DNA primer–template was described by Arnold and co-workers [40]. The nucleic acid is bound to the large polymerase cleft and stretches from the polymerase core to the RNaseH active site. The duplex deviates significantly from ideal A or B-form. The DNA conformation near the polymerase active site more closely resembles A-form, whereas after a pronounced kink 6–7 base pairs from the polymerase site, it is almost B-form.

A kink in the nucleic acid was also observed in the recently reported structure of HIV-1 RT co-crystallized with an RNA pseudoknot [4•]. This pseudoknot, a high-affinity aptamer for RT that was identified by SELEX [41], competitively inhibits the binding of primer–template DNA [42]. The RNA inhibitor is kinked by 60° from straight co-axial stacking of stems S1 and S2. The kink creates a curved shape that optimizes the extensive contacts between the RNA inhibitor and both subunits of the RT heterodimer (Figure 3). Through electrostatic interactions of stem S2 and loop L1 with several basic residues in both helix I of the p66 thumb and the p66 fingers subdomains, the RNA stabilizes a conformation of the polymerase in which the tip of the thumb and the fingers subdomains are in close contact. It is interesting to note that RNA stem S2, which is adjacent to the polymerase active site, does not interact with the β ribbon (residues 58–77) lining the fingers subdomain. The binding mode of the pseudoknot is more reminiscent of that observed in T7 DNAP, in which the template strand traverses the fingers. The pseudoknot-binding surface on the

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Figure 4

Schematic stereo drawing of the T7 RNAP–T7 lysozyme complex. The lysozyme mostly interacts with residues from the N-terminal domain (orange); the binding site is far from the polymerase active site, suggesting an indirect mechanism of inhibition. The overall structure of this enzyme is not reminiscent of those of other polymerases, although the palm and fingers domains display striking homology to *E. coli* DNA polymerase I (KF). The coloring scheme indicating subdomains and modules is as follows: N-terminal domain (orange), thumb (green), palm (red), palm insertion (yellow), fingers (dark blue), specificity loop or 'pinky' (black), extended foot (cyan) and lysozyme (magenta).



p66–p51 heterodimer partially overlaps with the binding surface for duplex DNA [40]. The RNA inhibitor has more contacts with the p51 subunit than the duplex DNA does and it may possibly overlap the binding site for tRNA Lys3.

DNA-dependent RNA polymerases

Bacteriophage T7 RNA polymerase (T7 RNAP) is a 98 kDa monomeric enzyme that catalyzes RNA transcription in a promoter-specific manner. A low resolution apoenzyme structure was reported by Sousa *et al.* in 1993 [43]. The outline of the T7 RNAP structure is irregular, but it displays typical features that are characteristic of the pol I family of polymerases. The thumb, palm and fingers subdomains are augmented by a long extended specificity loop ('pinky'), a palm insertion and an extended foot subdomain. Jeruzalmi and Steitz [5•] have recently reported the 2.8 Å crystal structure of the T7 RNAP–T7 lysozyme complex (Figure 4). *In vivo*, this complex controls gene expression during T7 phage infection of *E. coli*. A large body of biochemical and genetic data on promoter–DNA interactions [44] and on T7 lysozyme inhibition [45,46] has allowed detailed structure and function analyses. The new T7 RNAP crystal structure has a similar overall shape to that of the apoenzyme previously presented by Sousa *et al.* ([43], PDB code 2rnp); however, the connectivity in the N-terminal domain and the secondary-structure assignments throughout the polymerase have been revised and redefined.

The N-terminal 325 residues of T7 RNAP have no counterpart in other polymerase structures. This domain forms the front wall of the polymerase cleft and appears to be involved in promoter binding. Mutations in this area are detrimental to processive RNA synthesis. In contrast to the DNA pol I subfamily, the tip of the T7 RNAP thumb does not contain any strictly conserved residues. It also differs in shape from those found in other polymerases, indicating that the thumb subdomain may play a different role in substrate binding. Processivity may be controlled by other mechanisms in T7 RNAP. The fingers subdomain, located at the opposite side of the active site cleft, has a considerably higher degree of sequence conservation (J Jäger, unpublished data) and a curved shape that is compatible with the duplex template. Jeruzalmi and Steitz [5•] present clear evidence that the organization of the thumb subdomain in T7 RNAP is reminiscent of the corresponding domain in KF and T7 DNAP. Therefore, the incoming duplex may pass the fingers in T7 RNAP and then the template strand may kink sharply in order to enter the polymerase active site. Confirmation of this, however, requires further studies of large DNA operator–polymerase complexes and the results of such co-crystal structures are eagerly awaited.

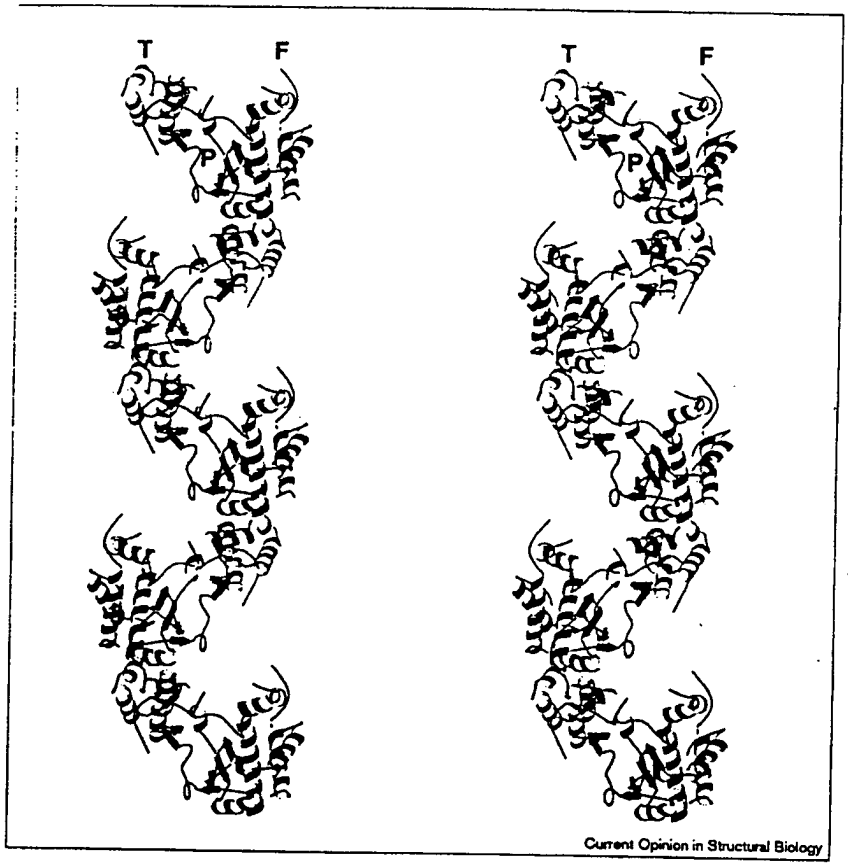
RNA-dependent RNA polymerases

Poliovirus 3D RNA polymerase (3Dpol) is the only RNA-dependent RNA polymerase whose three-dimensional structure has been determined [6•]. This class of polymerases is responsible for replicating and transcribing the genomes of nonretroviral RNA viruses. Even though the sequence similarity to other polymerase classes is very limited, the 3Dpol structure has the characteristic fingers, palm and thumb subdomains found in all polymerases, clearly indicating that the RNA-dependent RNA polymerases are evolutionarily related to the RTs and the DNA-dependent DNA and RNA polymerases. In particular, the core structure of the palm subdomain has the same topology as in the other polymerases and is composed of the four sequence motifs, A–D [27], that are common to all the polymerase families. The strictly conserved aspartate in motif A and the first of the two adjacent aspartates in the YGDD sequence of motif C are structurally equivalent to the two aspartates conserved in the other polymerase classes. Mutations of these aspartates abolish 3Dpol activity and mutation of the second aspartate in motif C alters the metal specificity of the enzyme [47].

The thumb and fingers subdomains of 3Dpol have little structural similarity to those in the other polymerases.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Figure 5



Stereo drawing of intermolecular contacts between poliovirus RNA polymerase molecules. Adjacent polymerase molecules (alternately shown in light and dark gray) pack in a head-to-tail manner, forming a fiber of polymerase molecules along a twofold screw axis in the crystal lattice. The interface between the front of the thumb of one molecule and the back of the palm of the next buries a surface area of 1480 Å^2 and involves at least 23 amino acid sidechains in a variety of hydrophobic, ionic and hydrogen bond interactions [7•]. Subdomains of the polymerase are identified as fingers (F), palm (P) and thumb (T).

although they presumably play similar roles in primer, template and NTP binding. The thumb subdomain is largely composed of sequences from the C terminus of the protein and is predominantly α helical. Helix K is positioned similarly to helix H in HIV-1 RT, which binds in the minor groove of the primer–template duplex [40]. Sequence motif E, common only to the RNA-dependent polymerases, is located between the palm and thumb subdomains. This motif forms a β-turn-β structure that was termed the 'primer grip' in HIV-1 RT [40] and that interacts with a β strand located at the beginning of the thumb subdomain in both 3Dpol and RT. Interestingly, this strand in the thumb of the poliovirus enzyme contains a valine residue that when mutated to isoleucine (V391I) specifically disrupts the interaction between the polymerase and the viral protein 3AB, both in the yeast two-hybrid system and *in vitro* [48]. 3AB is the precursor to the 22 amino acid protein VPg, which primes RNA synthesis by 3Dpol [49]. The fingers of 3Dpol are composed of sequences located N terminal to motif A and between motifs A and B. Only the lower portion of the fingers is ordered in the structure. Ribonucleotides cross-link to lysine residues in two peptide segments of 3Dpol (residues 266–286 and residues 57–74) [50] that are located adjacent to the ordered portions of the fingers. A detailed understanding of substrate binding by 3Dpol awaits future structure determinations.

The poliovirus polymerase crystal lattice reveals an unusual higher order structure that is believed to be important for polymerase activity. 3Dpol is known to bind and extend primer–template RNAs in a highly cooperative manner with respect to polymerase concentration [51], suggesting that the polymerase is active as an oligomer. Polymerase–polymerase interactions have been observed physically, by chemical cross-linking in solution [51], and genetically, in the yeast two-hybrid interaction assay [48]. In the crystals of 3Dpol, one interaction between polymerase molecules is an extensive interface between the front of the thumb of one molecule and the back of the palm of an adjacent molecule. This interaction produces a fiber of head-to-tail polymerase molecules (Figure 5) along a twofold screw axis in the crystal lattice and may correlate with the cooperative RNA-binding activity of 3Dpol. The binding site size for a single 3Dpol molecule on single-stranded RNA is approximately 10 nucleotides [52], a length that would be sufficient to span the 46 Å distance between two active sites along the polymerase fiber. The polymerase–polymerase interactions and the cooperative

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

RNA binding activity of poliovirus 3Dpol may not be unique. DNA pol β cooperatively binds substrate DNA, an activity that may serve to localize pol β to sites that require DNA repair [53–55].

Conclusions

The large database of polymerase structures currently available shows that all the polymerase domains have a common overall shape that resembles a right hand. As discussed above, the thumb, fingers and palm subdomains play similar functional roles, despite having minimal similarities in sequence and architecture. The similarities lead to a generalized image of polymerases, in which the primer–template duplex is bound by the thumb and positioned such that the 3′ end of the primer is located at the active site in the palm subdomain and the 5′ end of the template strand continues across the fingers subdomain. The incoming nucleotide is bound by both the fingers and palm subdomains, presumably entering through the cleft between the fingers and thumb, near the 3′ end of the primer. Even though the duplex does not pass through the cleft between the fingers and thumb, it is partially encircled by the thumb, with helices at the base of the thumb tracking the minor groove of the growing duplex and the tip of the thumb bending across the top of the duplex. Finally, the most universal feature of the polymerases is the two-metal-ion catalytic mechanism, being conserved throughout all polymerase classes.

The next few years promise to be exciting, as structural and functional polymerase studies continue to converge. It will be interesting to see how the structural concepts developed thus far apply to the processive multisubunit enzymes. Zhang and Darst [56] have taken one of the first steps in this direction, with the structure determination of the α subunit of *E. coli* RNA polymerase. Structural studies of initiation complexes will be another important step in understanding polymerase activity. The elucidation of these structures and mechanisms will be eagerly awaited.

Note added in proof

The recently published structure of HIV-1 RT complexed with dNTP and a covalently tethered primer–template provides a detailed view of the RT polymerase active site [58]. As seen in the bacteriophage T7 DNAP–DNA complex [1•], the fingers subdomain tightens around the substrates and is transversed by the template strand. This new structure is an important step towards understanding nucleoside drug resistance.

Acknowledgements

We thank Tom Steitz, Graham Cheetham and Steve Schultz for sharing unpublished data and for useful discussions. We gratefully acknowledge Jeff Hansen and Steve Smerdon for their critical reading of the manuscript and Jonathan Porter and Damien O'Farrell for help with the preparation of the manuscript. Figures were composed using the programs O, Ribbons, Bobscript and Spock. This work was partially supported by the Wellcome Trust and Yorkshire Cancer Research. JDP is funded by National Institutes of Health Fellowship number AI09695.

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. • Doublie S, Tabor S, Long AM, Richardson CC, Ellenberger T: **Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution.** *Nature* 1998, 391:252-258.
The authors describe a full quaternary complex of T7 DNA polymerase, with DNA, dideoxynucleotides and metal ions bound at the polymerase active site. This complex provides unprecedented insights into the mechanisms of polymerase inhibition, fidelity and processivity at high resolution.

2. • Kiefer JR, Mao C, Braman JC, Beese LS: **Visualizing DNA replication in a catalytically active *Bacillus* DNA polymerase crystal.** *Nature* 1998, 391:304-307.
This crystal form of a thermostable polymerase shows primer–template DNA bound productively at the active site. The 3′ end of the primer strand is not terminated and, thus, is available for catalytic turn-over. The resulting difference electron density maps show the incorporation of the correct nucleotides at near atomic resolution. The fingers subdomain is found in a different conformation from that described in [1•].

3. • Sawaya MR, Prasad R, Wilson SH, Kraut J, Pelletier H: **Crystal structures of human DNA polymerase beta complexed with gapped and nicked DNA: evidence for an induced fit mechanism.** *Biochemistry* 1997, 36:11205-11215
The authors describe three high-resolution X-ray structures of polymerase β complexed with gapped and nicked double-stranded DNA. The protein undergoes a large conformational change upon formation of the DNA complexes. The structures represent intermediates in the gap-filling reaction of polymerase β.

4. • Jaeger J, Restle T, Steitz TA: **The structure of HIV-1 reverse transcriptase complexed with an RNA pseudoknot inhibitor.** *EMBO J* 1998, 17:4535-4542.
This low resolution co-crystal structure shows how a RNA aptamer competitively inhibits HIV-1 reverse transcriptase (RT). The RNA ligand has been selected to bind to the closed, inactive form of RT.

5. • Jeruzalmi D, Steitz TA: **Structure of T7 RNA polymerase complexed to the transcriptional inhibitor T7 lysozyme.** *EMBO J* 1998, 17:4101-4113.
This co-crystal structure reveals that the T7 RNA polymerase inhibitor T7 lysozyme binds to a site that is distant from the polymerase active site cleft. The authors suggest an indirect mechanism of inhibition. For the first time, all of the subdomains of T7 polymerase are described in detail.

6. • Hansen JL, Long AM, Schultz SC: **Structure of the RNA-dependent RNA polymerase of poliovirus.** *Structure* 1997, 5:1109-1122.
The first glimpse of a RNA-dependent RNA polymerase. The fingers subdomain of the poliovirus polymerase is disordered, but a detailed analysis of the remaining structure highlights the evolutionary relationships to the other polymerase families. Curiously, the crystal packing suggests a mode of oligomerization that appears to be important for the enzyme's activity.

7. Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA: **Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP.** *Nature* 1985, 313:762-766.

8. Sakon J, Liao HH, Kanikula AM, Benning MM, Rayment I, Holden HM: **Molecular structure of kanamycin nucleotidyltransferase determined to 3.0 Å resolution.** *Biochemistry* 1993, 32:11977-11984.

9. Pedersen LC, Benning MM, Holden HM: **Structural investigation of the antibiotic and ATP-binding sites in kanamycin nucleotidyltransferase.** *Biochemistry* 1995, 34:13305-13311.

10. Holm L, Sander C: **DNA polymerase β belongs to an ancient nucleotidyltransferase superfamily.** *Trends Biochem Sci* 1995, 20:345-347.

11. Kornberg A, Baker TA: *DNA Replication*, edn 2. New York: WH Freeman; 1992.

12. Delarue M, Poch O, Tordo N, Moras D, Argos P: **An attempt to unify the structure of polymerases.** *Protein Eng* 1990, 3:461-467.

13. Joyce CM, Steitz TA: **Function and structure relationships in DNA polymerases.** *Annu Rev Biochem* 1994, 63:777-822.

14. Ito J, Braithwaite DK: **Compilation and alignment of DNA polymerase sequences.** *Nucleic Acids Res* 1991, 19:4045-4057.

15. Kim Y, Eom SH, Wang J, Lee DS, Suh SW, Steitz TA: **Crystal structure of *Thermus aquaticus* DNA polymerase.** *Nature* 1995, 376:612-616.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

16. Eom SH, Wang J, Steitz TA: **Structure of Taq polymerase with DNA at the polymerase active site.** *Nature* 1996, 382:278-281.

17. Kiefer JR, Mao C, Hansen CJ, Basehore SL, Hofgrefe HH, Braman JC, Beese LS: **Crystal structure of a thermostable *Bacillus* DNA polymerase I large fragment at 2.1 Å resolution.** *Structure* 1997, 5:95-108.

18. Wang J, Sattar AKMA, Wang CC, Karam JD, Konigsberg WH, Steitz TA: **Crystal structure of a pol α family replication DNA polymerase from bacteriophage RB69.** *Cell* 1997, 89:1087-1089.

19. Brautigam CA, Steitz TA: **Structural and functional insights provided by crystal structures of DNA polymerases and their substrate complexes.** *Curr Opin Struct Biol* 1998, 8:54-63.

20. Sawaya MR, Pelletier H, Kumar A, Wilson SH, Kraut J: **Crystal structure of rat DNA polymerase β: evidence for a common polymerase mechanism.** *Science* 1994, 264:1930-1935.

21. Georgiadis MM, Jessen SM, Ogata CM, Telesnitsky A, Goff SP, Hendrickson WA: **Mechanistic implications from the structure of a catalytic fragment of Moloney murine leukemia virus reverse transcriptase.** *Structure* 1995, 3:879-892.

22. Steitz TA: **DNA- and RNA-dependent DNA polymerases.** *Curr Opin Struct Biol* 1993, 3:31-38.

23. Steitz TA, Steitz JA: **A general two-metal-ion mechanism for catalytic RNA.** *Proc Natl Acad Sci USA* 1993, 90:6498-6502.

24. Brautigam CA, Steitz TA: **Structural principles for the inhibition of the 3'-5' exonuclease activity of *Escherichia coli* DNA polymerase I by phosphorothioates.** *J Mol Biol* 1998, 277:363-377.

25. Beese LS, Derbyshire V, Steitz TA: **Structure of DNA polymerase I Klenow fragment bound to duplex DNA.** *Science* 1993, 260:352-355.

26. Minnick DT, Astatke M, Joyce CM, Kunkel TA: **A thumb subdomain mutant of the large fragment of *Escherichia coli* DNA polymerase I with reduced DNA binding affinity, processivity, and frameshift fidelity.** *J Biol Chem* 1996, 24954-24961.

27. Poch O, Sauvaget I, Delarue M, Tordo N: **Identification of four conserved motifs among the RNA-dependent polymerase encoding elements.** *EMBO J* 1989, 8:3867-3874.

28. Steitz TA, Smerdon S, Jaeger J, Wang J, Kohlstaedt LA, Friedman JM, Beese LS, Rice PA: **Two DNA polymerases: HIV reverse transcriptase and the Klenow fragment of *Escherichia coli* DNA polymerase I.** *Cold Spring Harbor Symp Quant Biol* 1993, 58:495-504.

29. Davies JF, Almassy RJ, Hostomska Z, Ferre RA, Hostomsky Z: **2.3 Å crystal structure of the catalytic domain of DNA polymerase beta.** *Cell* 1994, 76:1123-1133.

30. Pelletier H, Sawaya MR, Kumar A, Wilson SH, Kraut J: **Structures of ternary complexes of rat DNA polymerase beta, a DNA template-primer and ddCTP.** *Science* 1994, 264:1891-1903.

31. Steitz TA, Smerdon SJ, Jaeger J, Joyce CM: **A unified polymerase mechanism for nonhomologous DNA and RNA polymerases.** *Science* 1994, 266:2022-2025.

32. Kohlstaedt LA, Wang J, Friedman JM, Rice PA, Steitz TA: **Crystal structure at 3.5 Å resolution of HIV-1 reverse transcriptase complexed with an inhibitor.** *Science* 1992, 256:1783-1790.

33. Smerdon SJ, Jaeger J, Wang J, Kohlstaedt LA, Chirino AJ, Friedman JM, Rice PA, Steitz TA: **Structure of the binding site for nonnucleoside inhibitors of the reverse transcriptase of human immunodeficiency virus type 1.** *Proc Natl Acad Sci USA* 1994, 91:3911-3915.

34. Rodgers DW, Gamblin SJ, Harris BA, Ray S, Culp JS, Hellmig B, Woolf DJ, Debouck C, Harrison SC: **The structure of unliganded reverse transcriptase from the human immunodeficiency virus type 1.** *Proc Natl Acad Sci USA* 1995, 92:1222-1226.

35. Unge T, Knight S, Bhikhabhai R, Lövgren, Dauter Z, Wilson K, Strandberg B: **2.2 Å resolution structure of the amino-terminal half of HIV-1 reverse transcriptase (fingers and palm subdomains).** *Structure* 1994, 2:953-961.

36. Jaeger J, Smerdon SJ, Wang J, Boisvert DC, Steitz TA: **Comparison of three different crystal forms shows HIV-1 reverse transcriptase displays an internal swivel motion.** *Structure* 1994, 2:869-876.

37. Das K, Ding J, Hsiou Y, Clark AD, Moereels H, Koymans L, Andries K, Pauwels R, Janssen PAJ, Boyer PL *et al.*: **Crystal structures of 8-Cl and 9-Cl TIBO complexed with wild-type HIV-1 RT and 8-Cl TIBO complexed with the Tyr181Cys HIV-1 RTdrug-resistant mutant.** *J Mol Biol* 1996, 264:1085-1100.

38. Ren J, Esnouf R, Garman E, Somers D, Ross C, Kirby I, Keeling J, Darby G, Jones Y, Stuart D, Stammers D: **High resolution structures of HIV-1 RT from four RT-inhibitor complexes.** *Nat Struct Biol* 1995, 2:293-302.

39. Esnouf R, Ren J, Ross C, Jones Y, Stammers D, Stuart D: **Mechanism of inhibition of HIV-1 reverse transcriptase by non-nucleoside inhibitors.** *Nat Struct Biol* 1995, 2:303-308.

40. Jacobo-Molina A, Ding J, Nanni RG, Clark AD, Lu X, Tantillo C, Williams RL, Kamer G, Ferris AL, Arnold E *et al.*: **Crystal structure of human immunodeficiency virus type 1 reverse transcriptase complexed with double-stranded DNA at 3.0 Å resolution shows bent DNA.** *Proc Natl Acad Sci USA* 1993, 90:6320-6324.

41. Tuerk C, MacDougal S, Gold L: **RNA pseudoknots that inhibit human immunodeficiency virus type 1 reverse transcriptase.** *Proc Natl Acad Sci USA* 1992, 89:6988-6992.

42. Burke DH, Scates L, Andrews K, Gold L: **Bent pseudoknots and novel RNA inhibitors of type 1 human immunodeficiency virus (HIV-1) reverse transcriptase.** *J Mol Biol* 1996, 264:650-666.

43. Sousa R, Chung YJ, Wang BC: **Crystal structure of bacteriophage T7 RNA polymerase at 3.3 Å resolution.** *Nature* 1993, 364:593-599.

44. Ujvari A, Martin CT: **Identification of a minimal binding element within the T7 RNA polymerase promoter.** *J Mol Biol* 1997, 273:775-781.

45. Kumar A, Patel SS: **Inhibition of T7 RNA polymerase: transcription initiation and transition to elongation are inhibited by T7 lysozyme via a ternary complex with RNA polymerase and promoter DNA.** *Biochemistry* 1997, 36:13945-13962.

46. Zhang X, Studier FW: **Mechanism of inhibition of bacteriophage T7 RNA polymerase by T7 lysozyme.** *J Mol Biol* 1997, 269:10-27.

47. Jablonski SA, Luo M, Morrow CD: **Enzymatic activity of poliovirus RNA polymerase mutants with single amino acid changes in the conserved YGDD amino acid motif.** *J Virol* 1991, 65:4565-4572.

48. Hope DA, Diamond SE, Kirkegaard K: **Genetic dissection of interactions between poliovirus 3D pol and viral protein 3AB.** *J Virol* 1997, 71:9490-9498.

49. Paul AV, Boom JV, Filipov D, Wimmer E: **Protein-primed RNA synthesis by purified poliovirus RNA polymerase.** *Nature* 1998, 393:280-284.

50. Richards OC, Hansen JL, Schultz SC, Ehrenfeld E: **Identification of nucleotide binding sites in the poliovirus RNA polymerase.** *Biochemistry* 1995, 34:6288-6295.

51. Pata JD, Schultz SC, Kirkegaard K: **Functional oligomerization of poliovirus RNA-dependent RNA polymerase.** *RNA* 1995, 1:466-477.

52. Beckman MTL, Kirkegaard K: **Site size of cooperative single-stranded RNA binding by poliovirus RNA-dependent RNA polymerase.** *J Biol Chem* 1997, 273:6724-6730.

53. Kumar A, Widen SG, Williams KR, Kedar P, Karpel RL, Wilson SH: **Studies of the domain structure of mammalian DNA polymerase beta. Identification of a discrete template binding domain.** *J Biol Chem* 1990, 265:2124-2131.

54. Casas-Finet JR, Kumar A, Morris G, Wilson SH, Karpel RL: **Spectroscopic studies of the structural domains of mammalian DNA beta-polymerase.** *J Biol Chem* 1991, 266:19618-19625.

55. Casas-Finet JR, Kumar A, Karpel RL, Wilson SH: **Mammalian DNA polymerase beta: characterization of a 16-kDa transdomain fragment containing the nucleic acid-binding activities of the native enzyme.** *Biochemistry* 1992, 31:10272-10280.

56. Zhang G, Darst SA: **Structure of the *Escherichia coli* RNA polymerase α subunit amino-terminal domain.** *Science* 1998, 281:262-266.

57. Li Y, Kong Y, Korolev S, Waksman G: **Crystal structures of the Klenow fragment of *Thermus aquaticus* DNA polymerase I complexed with deoxyribonucleoside triphosphates.** *Protein Sci* 1998, 7:1116-1123.

58. Huang H, Chopra R, Verdine GL, Harrison SC: **Structure of a covalently trapped catalytic complex of HIV-1 reverse transcriptase: implications for drug resistance.** *Science* 1998, 282:1669-1675.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Institut für Molekulare Biotechnologie JENA

Peter Slickers (slickers@imb-jena.de)
Institut für Molekulare Biotechnologie, Beutenbergstraße 11, D-07745 Jena, Germany
Phone: +49-3641-6202, Fax: +49-3641-6210

DNA Polymerases

Abstract

To date the three-dimensional structures of 8 different nucleic acid polymerases or their subunits are known. Similar structures are found for all the domain which mediate the polymerase or transphosphorylation activity. All this domains have a hand-like structure with a palm, fingers and thumb subdomain. Three carboxylic side chains, which are always found in the same spatial arrangement, and two magnesium ions are directly involved in the transphosphorylation reaction. Two aspartatic acids are oriented nearly parallel to each other and coordinate two magnesium ions. The magnesium ions are also in close contact to phosphate and hydroxy groups of the primer and the incoming mononucleotide. One magnesium is thought to stabilize the transition state of the transphosphorylation. This two metal-ion mechanism is unique to polymerase. High processivity is mediated to the core of a polymerase by a ring-shaped protein complex. This processivity factor act as DNA clamps. The processivity factors are highly symmetrical and consist of two or three identical monomers.

Content

1. Introduction

Nucleic acid polymerases are enzymes found in all living cell. DNA polymerases are used the replicating of DNA, while RNA polymerases do the transcription of DNA into RNA. Further on, some DNA polymerases are able to repair DNA. All the polymerases involved in replication, transcription and repair of

DNA are template-directed. They need a single strand of DNA as a template to synthesize a second strand of nucleic acid which is complementary in sequence to the template strand. DNA polymerases add deoxynucleotides to the growing strand, while RNA polymerases add ribonucleotides. Only some special polymerases like the telomerases do not need a template to synthesize new DNA.

In general, nucleic acid polymerases are catalysing the formation of a bond between a polynucleotide chain, which is called the primer, and a mononucleotide. The template strand is a cofactor for this reaction and also two magnesium ions as will be further discussed bellow. More precisely, the polymerisation reaction is a transphosphorylation. This is illustrated in Figure 1. The educts are a primer with a free 3' hydroxy group and a mononucleoside-triphosphate. The 3' oxygen makes a nucleophilic attack on the aplha phosphate (the innermost phosphate) of the mononucleotide and the beta and gamma phosphates are released as pyrophosphate. (So a nucleotide-monophosphate is transphosphorylated from a pyrophosphate to the 3' oxygen of the primer.) The transphosphorylation is energetically unfavourable, but the reaction is driven forward by the immediate hydrolysation of the pyrophosphate into two anorganic phosphates. This is catalysed by a pyrophosphatase.

Figure 1 also illustrates that all template-dependend nucleic acid polymerases read the template strand in 3' to 5' direction and add mononucleotides with their 5' phosphate to the 3' end of the growing polynucleotide chain. Therefore the reaction is also called a 5'-3' polymerisation. RNA polymerases are able to use a single mononucleotide as the primer, whereas DNA polymerases always need an oligonucleotide. Accordingly only RNA polymerases are able to synthesize a polynucleotide *de novo*. But DNA polymerases are able to add deoxy-mononucleotides either to a DNA or a RNA primer. Consequently DNA polymerases use RNA primers to initialize new DNA strands.

According to their templates and their products, four types of template-dependend polymerases are known:

DNA-dependent DNA polymerase
: are employed in all cells for replication and repair of the DNA.

RNA-dependent DNA polymerase
: are found with retro viruses which use to copy RNA into DNA. This type of DNA polymerase are called reverse transcriptase (RT).

DNA-dependent RNA polymerase
: are needed in all cells for the transcription of DNA. Primases are also RNA polymerases and they are responsible for the synthesis of the RNA primers needed for the replication of the DNA.

RNA-dependent RNA polymerase
: are found with such RNA viruses which replicate by copying their RNA directly into new RNA.

Before a cell is ready to devide into two daughter cells, the genom has to be duplicated. The replication of the DNA double helix is a complex process and several proteins are involved. Replication starts with the unwinding of the double helix and the formation of a replication fork. While the replication fork moves along the parent DNA, one new strand (the leading strand) is synthesized continously. The other strand (the lagging strand) is synthesized discountinously. As DNA polymerase read their template strands only in 3' to 5' direction, one template strand has to be read against the direction of movement of the replication fork. This strand is the template for the lagging strand. The lagging strand is synthesized within fragments which are called Okazaki fragments and are 1000 to 2000 bases in length.

The synthesis of each Okasaki fragment requires a primer. In *E. coli* the primers are RNA oligonucleotides which are synthesized by a special RNA polymerase called DNA primase. Then the primers are extended by such DNA polymerases which have a high processivity and a high chain elongation rate. The RNA primers of the Okazaki fragments are replaced with DNA by another type of DNA polymerases, which have an additional 5'-3' exonuclease activity to cut of the primer. Finally a DNA ligase joins the Okazaki fragments covalently.

Thus there are several different polymerases involved in the replication of DNA *in vivo*. They do not only differ in their template and substrate specificity, but also in their processivity, their fidelity and their chain elongation rates. The processivity is the number of nucleotides a polymerase add to the growing strand before the polymerase dissociates from the template. The fidelity is a measure of the accuracy and describes how often a wrong nucleotide is incorporated. To keep the genetic information for the next generation a very high fidelity is required for the DNA replication. This is achieved by a proof-reading activity present with some DNA polymerases. The proof-reading activity is a 3'-5' exonuclease activity which removes mismatched nucleotides.

Most organisms have several different DNA polymerases. *E. coli* for example has three different DNA polymerases called polymerase I (Pol I), polymerase II (Pol II) and polymerase III (Pol III). Pol III has a high processivity and is the functional enzyme of DNA replication. Pol I has a low processivity and is responsible for replacing the RNA primers, which were used to initialize the new DNA strands, with DNA and for DNA repair. Less is known about Pol II. Five different polymerases are known for eukariotic cells, named alpha, beta, gamma, delta and epsilon. Presumably polymerase alpha, delta and epsilon are involved in the replication of the nuclear DNA, whereas polymerase gamma is located in the mitochondria. Polymerase beta has no proof-reading activity and a very low processivity and is thought to act in DNA repair.

---

2. The known 3-dimensional structures of polymerases: An overview

As outlined in the introduction cells are using several different polymerases for the tasks of DNA replication, transcription and repair. All this enzymes have been characterized biochemically and their genes are known. But less is know about their 3-dimensional structure. All known structures are deposited in the Brookhaven Protein Data Bank (PDB). To get a survey of the know structures, the PDB was searched (PDB Browse) for all polymerases and transcriptases. Currently (october 1996) the PDB contains solved crystall structures of 8 different polymerases or their subunits. 30 entries were found in total because more than one structure was solved for some polymerases. All the entries found are listed in Table 1.

5 of the 8 structures include the polymerase active site and the structure of this 5 proteins are described in more detail in this section. The other 3 structures are a 3'-5' exonuclease domain and two processivity factors. The structure of the processivity factors is further discussed in Section 4.

The first 3-dimensional structure of a nucleic acid polymerase to be solved was the structure of the Klenow fragment of Pol I of *E. coli* (Ollis et al. 1987 a). Pol I is a monomer with 928 amino acids and has three enzymatic activities: 5'-3' DNA polymerase, 3'-5' exonuclease (proof-reading) and 5'-3' exonuclease. The 5'-3' exonuclease acts in digesting RNA primers and in creating gaps during DNA repair which are refilled by the polymerase activity. Proteolytic cleavage of Pol I yields two fragments. The larger C-terminal fragment (Res 327-928) is called Klenow fragment and is a functional enzyme with 5'-3' polymerase and

3'-5' exonuclease (proof-reading) activity.

The second structure solved was reverse transcriptase (RT) from human immuno-deficiency virus (HIV) (Kohlstaedt et al. 1992). RT is a heterodimer of the two proteins P66 (66 kD) and P51 (51 kD). Both proteins are encoded by the same gene. After translation a 14 kD C-terminal domain is cleaved from P66 which yields P51. The C-terminal domain of P66 comprises a Ribonuclease H activity (5'-3' exonuclease). P66 also comprises the active polymerase site, whereas P51 has no enzymatic activity. This is surprising since both proteins have the same sequence. But the crystall structure reveals, that they fold into different spatial arrangements, although they have the same sequence and the same secondary structure elements. The tertiary structure of P51 is enzymatically inactive.

The overall shape of P66 is strikingly similar to Klenow fragment (KF). Both proteins are resembling a hand which grips around a rod. Consequently the subdomains were named according to the anatomy of a hand. The palm domain is located at the bottom of the cleft which is flanked by the fingers on one side and the thumb on the other. Later on this overall hand-like shape was also found for two other polymerases: bacteriophage T7 RNA polymerase (Sousa et al. 1993) and rat polymerase beta (Pelletier et al. 1994). The structure of the reverse transcriptase of Moloney murine leukemia virus (MMLV) is the fifth known structure of a polymerase (Georgiadis et al. 1995). But only a fragment of the enzyme was crystallized, and this fragment contains only the fingers and a palm domains, while the thumb is missing.

The hand-like shape common to all polymerases is illustrated in Figure 2 for the four polymerases which have a complete hand. Although the hand-like shape is found with all polymerases their sequence-similarity and their similarity in secondary structure is low. It is still an open question, whether the hand-like shape is a result of coevolution or a remaining of a common ancestor enzyme. Sousa (1996) argues for the existance of a superfamily of all polymerases. But the low sequence similarity and the great differences in the topology of the secondary structure elements found for the hand-like domain are more indicative of a homologous evolution, where the hand-like shape is a functional requirement. Little is known about the mechanical details of the translocation of the polymerases along the template. But it is obvious that the translocation must be precisely coupled to the transphosphorylation reaction and requieres a conformational change of the protein. Maybe the organisation in distinct palm, fingers and thumb domains, which are flexible against each other, is functionally related to the mechanism of translocation.

Another feature found with all polymerases is the location of three acidic amino acids at the polymerase active site. They are listed in Table 2 for the five polymerases with known structure. The location and arrangement of the three side chains can be seen from the RasMol views included in Figure 2. The three carboxylates are directly involved in the transphosphorylation reaction as described in detail in Section 3. The active site with the three carboxylic side chains is located at the bottom of the cleft between the fingers and thumb domains and on the surface of the palm domain. Again the question remains open whether the location of the three carboxylic side chains is a result of analogous or homologous evolution.

---

3. The mechanism of transphosphorylation

The structure of rat polymerase beta (Pelletier et al. 1994) allows the best insight into the mechanism of the transphosphorylation reaction. Pol beta is a very simple polymerase with only a 5'-3' polymerase activity, but no exonuclease or proof-reading activity. The enzyme is a monomer of 335 residues and was crystallized in complex with a template DNA strand, a primer DNA strand and a monomeric 2',3'-dideoxycytisine-triphosphate (ddCTP). In addition the 3' end of the primer was terminated by a 2',3'-dideoxycytosine. The missing 3' hydroxy group of the primer does not allow a transphosphorylation and freezes the complex in the initial state of the reaction.

The arrangement at the active site is shown in Figure 3. The primer as well as the incoming nucleotide are forming classical Watson-Crick hydrogen bonds with the template. The cytosine base of the incoming ddCTP stacks on the 3' cytosine of the primer. The 3' hydroxyl group of the primer is not present in the crystal structure, but was add in order to illustrate the reaction mechanism. The O3' of the primer nucleophilically attacks the alpha phosphorous of the mononucleotide. This leads to a transition state where the alpha phosphorous is penta-coordinated to five oxygens. Initially the structure of the alpha phosphate is tetrahedral, but changes to planar-bipyramidal in the transition state. The transition state is stabilized by the left magnesium ion in Figure 3, whereas the right magnesium stabilizes the leaving pyrophosphate and also helps in positioning the incoming mononucleotide. The three aspartatic acid residues Asp 190, Asp 192 and Asp 256 are crucial for placing the two magnesium ions, but they are not involved directly in catalysis.

---

4. DNA clamps mediate high processivity

The intact holoenzyme of *E. coli* Pol III is a large assembly of 10 different protein subunits (alpha, beta, gamma, delta, delta', epsilon, theta, tau, chi, psi). The polymerase activity is located on subunit alpha, and the 3'-5' exonuclease activity (proof-reading) on subunit epsilon. An assemly of alpha, epsilon and theta is called the core polymerase complex and can perform DNA polymerisation, but with a low processivity. The complex dissoziates from the template after adding a few nucleotides to the growing chain. The complex gamma, consisting of the gamma, delta, delta', chi and psi subunit, catalysizes to linkeage of two beta subunits to DNA and to the core complex. After the beta subunit is added to the core complex, the high processivity of the holoenzyme is recovered. Thus the beta subunit has no catalytic activity by itself, but mediates the high processivity to the core complex.

The crystall structure of the beta subunit was solved by Kong et al. 1992 and deposited in the PDB (2pol). The protein is a dimer of two identical subunits. Each beta subunit has 366 amino acid residues. The shape of the dimer is shown in Figure 4. The two monomers stick together head by tail and form a closed ring which wraps around a central pore. The ring has an overall diameter of 80 Å, and the diameter of the pore is approximately 35 Å. In the model shown in Figure 4 a DNA double helix is added to the experimentally derived protein structure to illustrate the presumable function of the beta subunit. The two monomers clamp around the DNA and, when linked to the core complex, they prevent the dissoziation of the polymerase/DNA complex.

The tertiary structure of the beta subunits is highly symmetrical (Figure 5) and therefore quite beautiful. 48 beta sheets are forming the outer surface, while 12 alpha helices are lining the inner surface. Each monomer consists of three domains and all of the three domains within one protein have the same secondary structure and nearly the same spacial arrangement. This is amazing because the amino acid sequences of the three domains are quite different. The structure of one such domain and its orientation towards the DNA is shown in Figure 6. The domain consists of 8 beta sheets and 2 alpha helices. Either four beta sheets are arranged anti-parallel in a Greek Key motif. The domain is twofold symmetrical by itself. The sheets and helices are enclosing a core of closely packed side chains. This core stabilizes the hole domain. By assembling 6 such building blocks a rigid ring is constructed. The beta subunit does not recognize certain DNA sequences. The pore of the ring is wide enough to avoid contacts between protein side chains and DNA bases. Only the sugar-phosphate backbone is recognized by the beta subunit. The task of the beta subunit is to duct the DNA to or from the core of the polymerase complex regardless of the DNA sequence.

Eukariotic cells have a processivity factor very similar to the beta subunit of *E. coli*. The protein is called proliferating cell nuclear antigene (PCNA). The tertiary structure Figure 7 is nearly identical to the beta

subunit, it also forms a ring enclosing the DNA helix. But with PCNA the ring consists of three rather than two monomers. The three monomers are identical. The function is the same for PCNA and the beta subunit. Both are DNA clamps mediating high processivity.

---

5. References

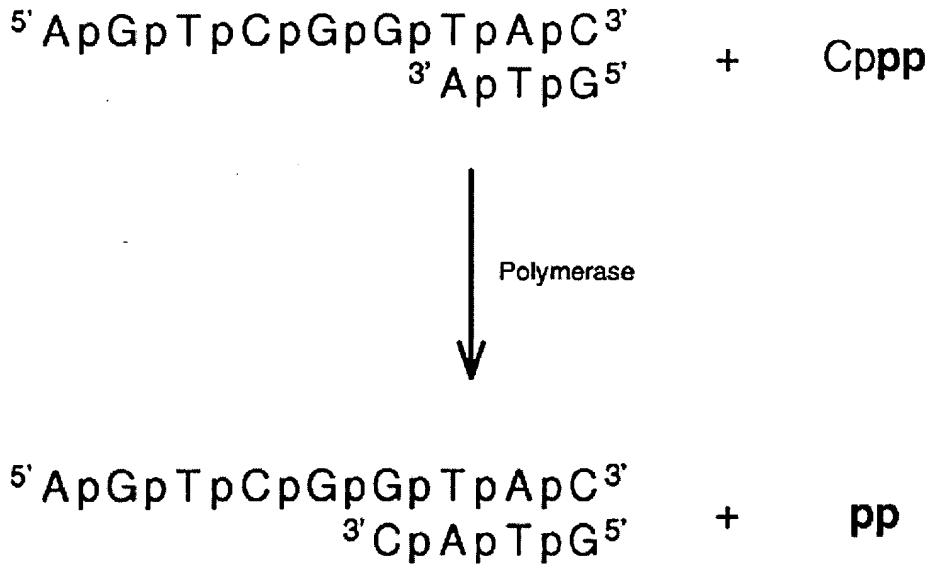
Table of Content

Top of this document

Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

1. Transphosphorylation



2. Hydrolysis of pyrophosphate

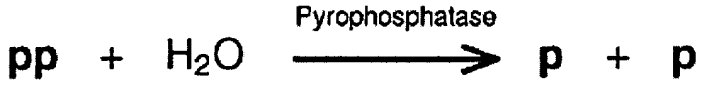


Figure 1.

The polymerisation reaction catalysed by DNA and RNA polymerases is a transphosphorylation. A mononucleoside-triphosphate is added to the 3' OH group of the growing chain by forming an phospho-ester bond between the alpha phosphate of the mononucleotide and the free 3' OH of the growing strand. So the alpha phosphate is incorporated in the backbone of the new polynucleotide chain, while the beta and gamma phophates are released as pyrophosphate. The transphosphorylation is energetically unfavourable, but the reaction is driven forward by the immediate hydrolysation of the pyrophosphate into two anorganic phosphates by a pyrophosphatase.

Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

Table 1.

Currently (october 1996) crystal structures of subunits or holoenzymes of 8 different proteins related to polynucleotide polymerases have been solved and deposited in the Brookhaven Protein Data Bank (PDB):

1. DNA polymerase beta from Rat
2. Klenow fragment of DNA polymerase I from *E. coli*
3. RNA polymerase from Bacteriophage T7
4. Reverse transcriptase from HIV-1
5. Catalytic fragment of reverse transcriptase from MMLV
6. 3'-5' Exonuclease domain of DNA polymerase from Bacteriophage T4
7. Beta subunit (processivity factor) of DNA polymerase III from *E. coli*
8. PCNA (processivity factor) of eukaryotic DNA polymerases from Yeast

1. DNA polymerase beta (Rat)

Protein	Ligand	Resolution	Reference	PDB code
31 KD catalytic domain (Res 91-335)		2.3 Å	Sawaya et al. 1994	1bpb
31 KD catalytic domain (Res 95-335)		2.3 Å	Davies et al. 1994	1rpl
31 KD catalytic domain (Res 91-335)	Mn^{2+}	2.8 Å	Sawaya et al. 1994	2bpc
39 KD (Res 9-335)	PO_4	3.6 Å	Sawaya et al. 1994	1bpd
39 KD (Res 12-335)	dATP	3.9 Å	Sawaya et al.	1bpe

			1994	
39 KD (Res 9-335)	ddCTP, template DNA, primer DNA, 2 Mg^{2+}	2.9 Å	Pelletier et al. 1994	2bpf
39 KD (Res 9-335)	ddCTP, template DNA, primer DNA	3.6 Å	Pelletier et al. 1994	2bpg

---

2. Klenow fragment of DNA polymerase I (*E. coli*)

Protein	Ligand	Resolution	Reference	PDB code
Klenow fragment (Res 326-928)	Zn^{2+}	2.8 Å	Ollis et al. 1985 a	1dpi (C alpha coordinates only)
Klenow fragment (Res 326-928)	dCTP	3.9 Å	Beese et al. 1993 b	1kfd
Klenow fragment mutante (Res 326-928)	DNA (Res 2-29), Zn^{2+}	3.2 Å	Beese et al. 1994 a	1kln

---

3. RNA polymerase (Bacteriophage T7)

Protein	Ligand	Resolution	Reference	PDB code
Res 1-883	Zn^{2+}, Mn^{2+}	3.3 Å	Sousa et al. 1993	2rnp (C alpha coordinates only)

---

4. Reverse transcriptase (HIV-1)

Protein	Ligand	Resolution	Reference	PDB code
fingers and palm (Res 1-216)	Zn^{2+}	2.2 Å	Unge et al. 1994	1har
ribonuclease H domain (Res 427-556)	Zn^{2+}	2.4 Å	Davies et al. 1991	1hrh
ribonuclease H			Chattopadhyay	

domain (Res 427-556)	Zn^{2+}	2.8 Å	et al. 1993	1rdh
P66 (Res 1-556), P51 (Res 2-427)	DNA template, DNA primer	3.0 Å	Jacobo-Molina et al. 1993	1hmi (C alpha coordinates only)
P66 (Res 1-554), P51 (Res 5-427)	Mg^{2+}	3.2 Å	Rodgers et al. 1994	1hmv
P66 (Res 1-558), P51 (Res 1-427)	alpha-APA	2.8 Å	Ding et al. 1995 a	1hni
P66 (Res 1-558), P51 (Res 1-427)	TIBO	3.0 Å	Ding et al. 1995 b	1hnv
P66 (Res 2-556), P51 (Res 2-428)	Neveripin	2.9 Å	Wang et al. 1994	3hvt
P66 (Res 1-543), P51 (Res 2-440)		2.4 Å	Esnouf et al. 1995	1rtj
P66 (Res 1-543), P51 (Res 5-437)	Benzodiazepinone	2.2 Å	Ren et al. 1995 a	1rth
P66 (Res 1-543), P51 (Res 3-438)	Phenyl-thiothymine	3.0 Å	Ren et al. 1995 a	1rti
P66 (Res 4-539), P51 (Res 5-428)	Nevirapine, Mg^{2+}	2.2 Å	Ren et al. 1995 a	1vrt
P66 (Res 3-539), P51 (Res 4-428)	alpha-APA	2.4 Å	Ren et al. 1995 a	1vru

---

5. Catalytic fragment of reverse transcriptase (MMLV)

Protein	Ligand	Resolution	Reference	PDB code
catalytic fragment (Res 24-274)		1.8 Å	Georgiadis et al. 1995	1mml

---

6. 3'-5' Exonuclease domain of DNA polymerase (Bacteriophage T4)

Protein	Ligand	Resolution	Reference	PDB code
Res 2-370	Zn^{2+}, Mn^{2+}	2.2 Å (298 K)	Wang et al. 1996	1noy
Res 2-370		2.2 Å (110 K)	Wang et al. 1996	1noz

---

7. Beta subunit (processivity factor) of DNA polymerase III (*E. coli*)

Protein	Ligand	Resolution	Reference	PDB code
2 monomers of identical sequence (Res 1-366)		2.5 Å	Kong et al. 1992	2pol

---

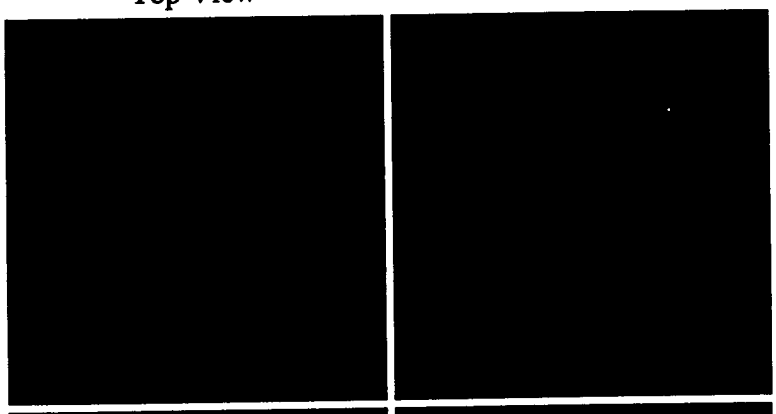
8. PCNA (processivity factor) of eukaryotic DNA polymerases (Yeast)

Protein	Ligand	Resolution	Reference	PDB code
1 monomer (Res 1-258)	2 Hg^{2+}	2.3 Å (synchrotron)	Krishna et al. 1994	1plq
1 monomer (Res 1-258)		3.0 Å	Krishna et al. 1994	1plr

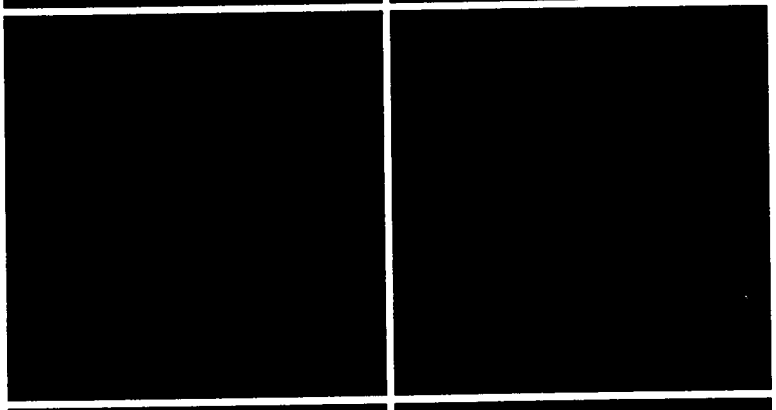
Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

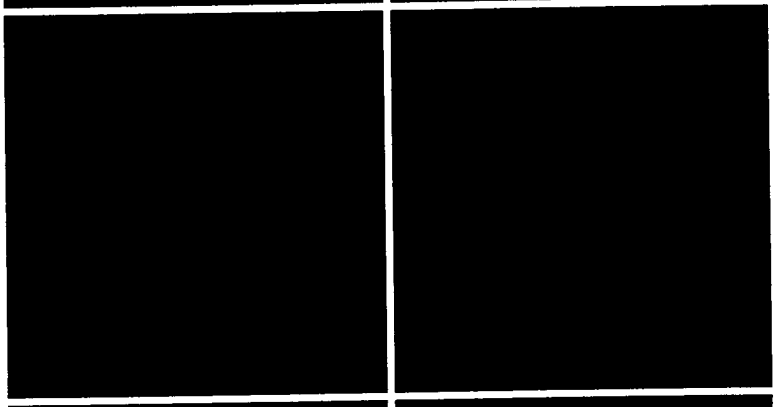
Top View | Side View



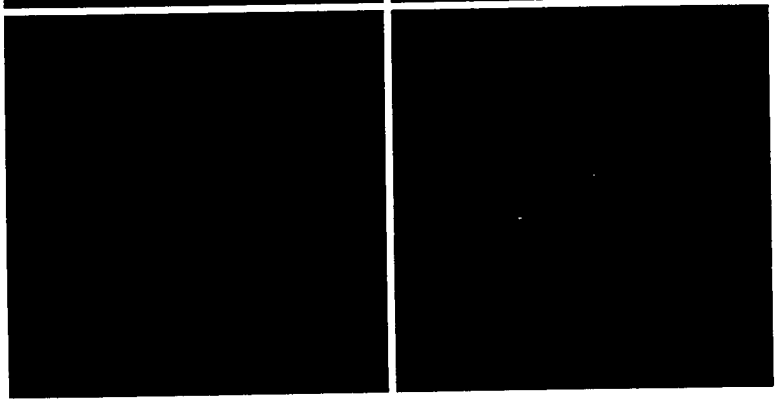
P66 reverse transcriptase (HIV)
(RasMol script)



Klenow fragment
(RasMol script)



Pol beta (Rat)
(RasMol script)



RNA Pol (Bacteriophage T7)
(RasMol script)

Figure 2.

The domain containing the polymerase active site has a hand-like shape with fingers (**blue**), palm (**magenta**) and thumb (**green**). The hand-like shape was found for all polymerases for which the structure of the active site was solved. In addition the three acidic side chains, which are found in the same spatial arrangement in all polymerases, are shown. They are located at the bottom of the cleft (palm domain) and are essential for catalysis. The RasMol scripts allow to have a more detailed look for this. The distances between the three carboxyl groups are shown within the RasMol views (reference atoms are C gamma for Asp and C delta for Glu). The orientation of the side chains of T7 RNA polymerase is tentative since the coordinates of the side chains are not included in the PDB entry.

Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

Overview (RasMol script)

Zoom (RasMol script)

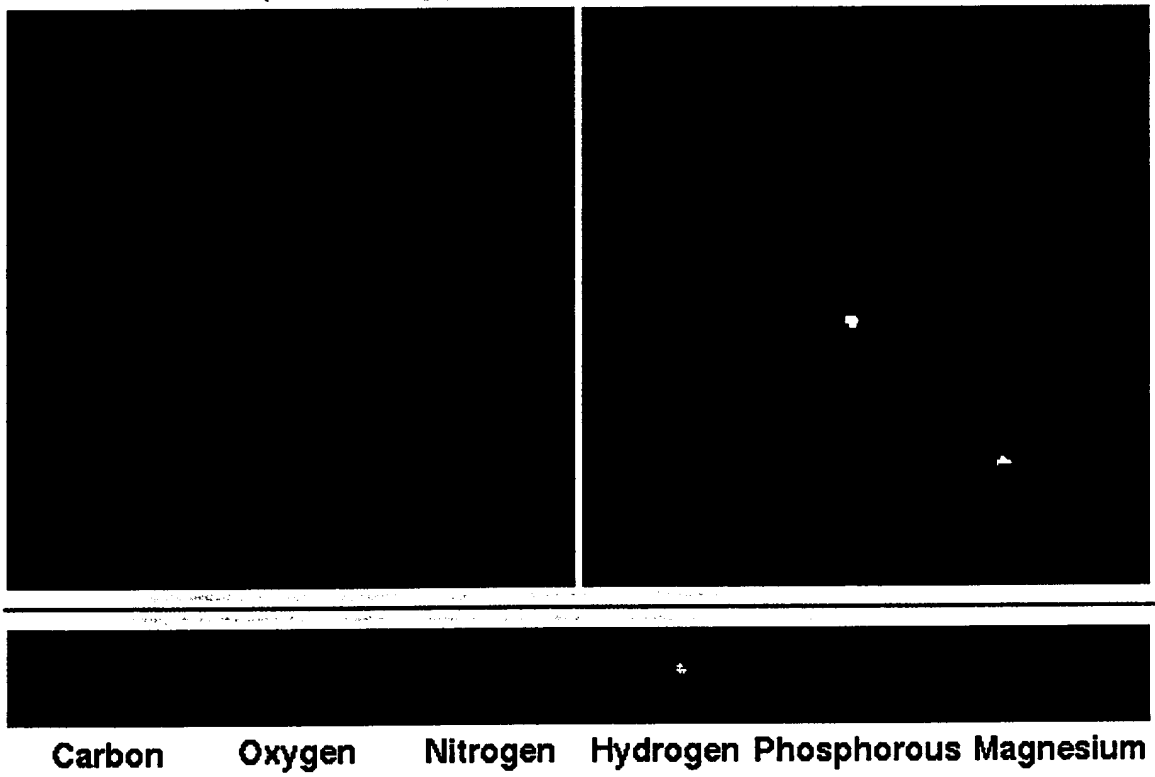


Figure 3.

The active site of rat Pol beta (PDB entry 2bpf). The DNA template, the DNA primer, the incoming mononucleotide dCTP, three aspartatic acid side chains and two magnesium ions are shown. The 3' hydroxyl group of the primer is not present in the crystal structure, but was add to illustrate the reaction mechanism.

The primer as well as the incoming nucleotide are forming classical Watson-Crick hydrogen bonds to the template. The oxygen O3' of the primer nucleophilically attacks the alpha phosphorous of the mononucleotide. The transition state of this is stabilized by the left magnesium ion, whereas the right magnesium helps to position the incoming mononucleotide and stabilizes the leaving pyrophosphate.

Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

Front View (RasMol script) Side View (RasMol script)

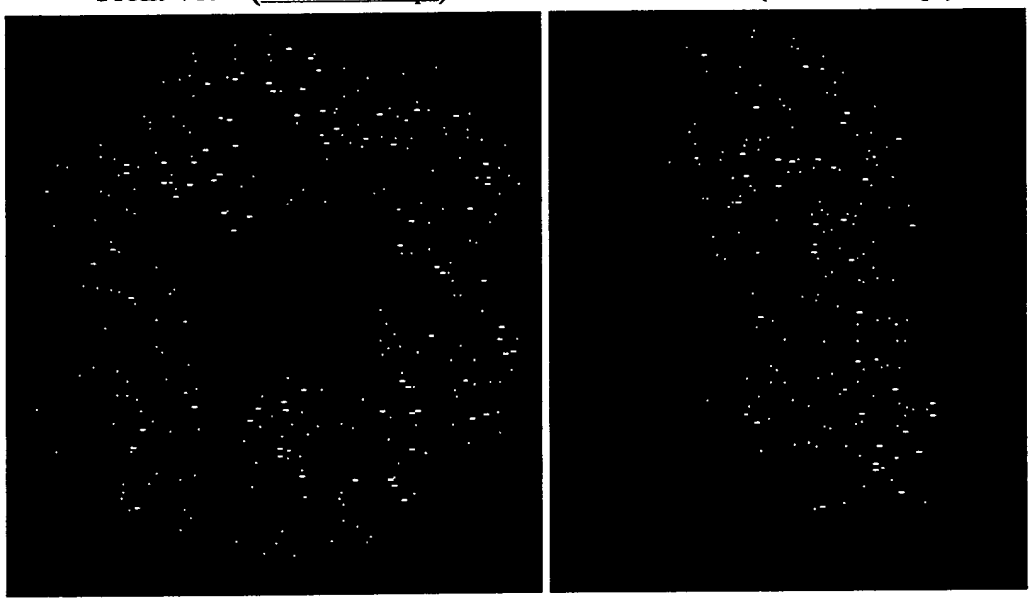


Figure 4.

Two beta subunits (red and blue) of *E. coli* Pol III form a ring which wraps around the DNA. The coordinates of the protein were taken from PDB entry 2pol (Kong et al. 1992). The DNA is an ideal B-type double helix, which was placed hypothetically to pass through the pore of the protein dimer.

Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

Front View (RasMol script)

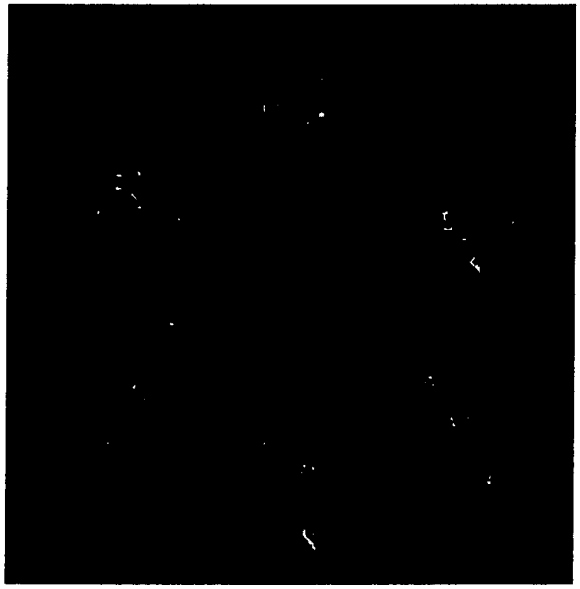


Figure 5.

The tertiary structure of the beta subunit of Pol III is highly symmetrical. The ring shape of the protein is reflected by the arrangement of the 48 beta sheet and the 12 alpha helices. The alpha helices are lining the inner surface, while the beta sheets are forming the outer surface.

Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

Front View (RasMol script) Side View (RasMol script)

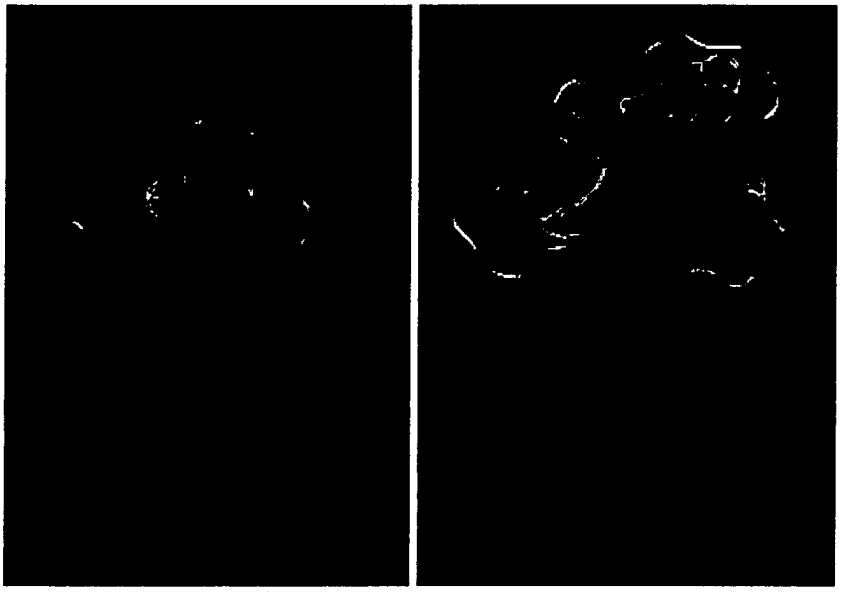


Figure 6.

Each beta subunit consists of three domains with identical secondary structure and nearly the same 3-dimensional arrangement. One such domain is shown here in a Richardson type schematic diagram. The domain consists of two alpha helices and eight beta sheets and is symmetrical by itself. As can be seen from the side view either four beta sheets are arranged anti-parallel in a Greek Key motif.

Peter Slickers (slickers@imb-jena.de)

DNA Polymerases

Front View (RasMol script) Side View

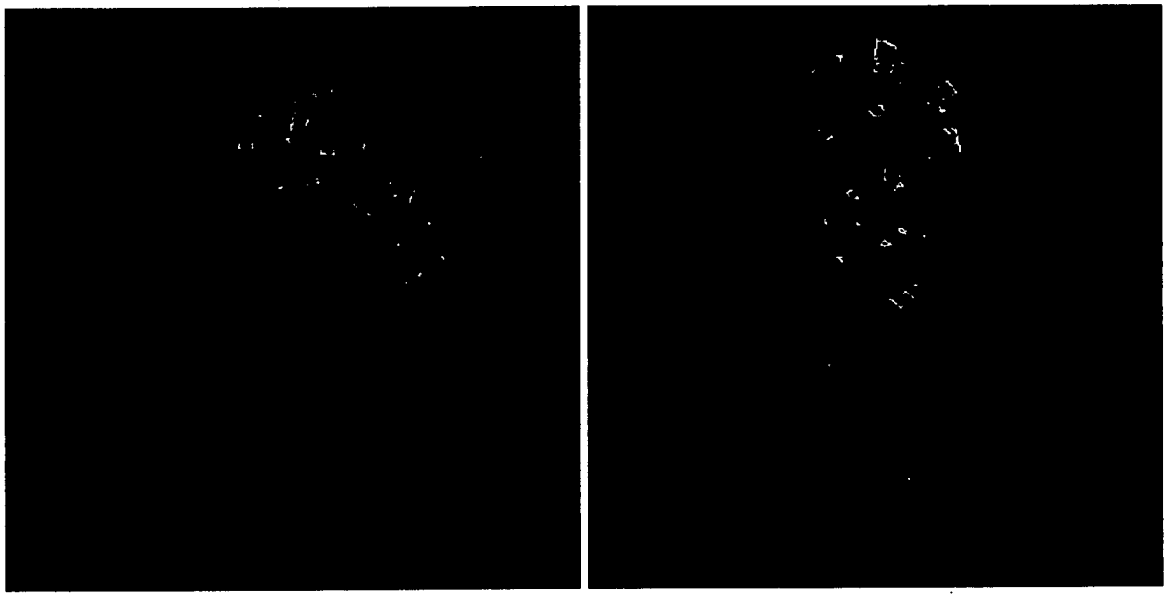


Figure 7.

The tertiary structure of the eukaryotic processivity factor PCNA (proliferating cell nuclear antigene). PCNA (PDB entry 1plq) is a trimer of three identical subunits shown in red, yellow and green. The three subunits form a ring-shaped structure. PCNA is a DNA clamp and mediates high processivity to the eukaryotic polymerases delta and epsilon. Although PCNA consists of three monomers the secondary and tertiary structure are very similar to subunit beta of *E. coli* Pol III.

TiBS

trends in BIOCHEMICAL SCIENCES

is an official publication of the INTERNATIONAL UNION OF BIOCHEMISTRY AND MOLECULAR BIOLOGY

EDITORIAL BOARD

Tim Hunt (editor-in-chief), Herts, UK
Joanne Chory, La Jolla, USA
Jack Dixon, University of Michigan, USA
Alan Fersht, Cambridge, UK
Mary-Jane Gething, Melbourne University, Australia
Carol Greider, New York, USA
Carol-Henrik Heldin, Uppsala, Sweden
Matthias Hentze, Heidelberg, Germany
Roger Kornberg, Stanford, USA
Angus Lamond, Dundee, UK
Thomas Pollard, La Jolla, USA
Paul Schimmel, Cambridge, USA
Peter H. von Hippel, Eugene, USA
Jan Witkowski, New York, USA
Mitsuhiro Yanagida, Kyoto, Japan

EDITOR — Jo McEntyre

Assistant Editor — Mark Swallow

Editorial Administrator — Christine Oliver and Gaynor Darke

Publisher — Peter Desmond

REFERENCE EDITION
Volume 21
1996

ELSEVIER TRENDS JOURNALS

Published by
The International Union of Biochemistry and Molecular Biology and Elsevier Trends Journals

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

Structural and mechanistic relationships between nucleic acid polymerases

Rui Sousa

A superfamily of nucleic acid polymerases that includes the pol I and pol α classes of DNA-directed DNA polymerases, mitochondrial and phage DNA-directed RNA polymerases, and most RNA-directed polymerases may be defined on the basis of the occurrence of conserved sequence motifs and tertiary structure similarities between HIV-1 reverse transcriptase, DNA polymerase I and T7 RNA polymerase. Although sequence or structural similarities do not yet justify inclusion of the multi-subunit DNA-directed RNA polymerases in this superfamily, mechanistic similarities suggest a deep relationship between these and the simpler T7-like RNA polymerases.

THE TERTIARY STRUCTURE of a template-directed nucleic acid polymerase, that of the Klenow fragment of DNA polymerase I (DNAP I), was first described in 1985 (Ref. 1). Seven years were to pass before another polymerase structure appeared in the literature. During this period, development of the field depended largely on structure–function studies and on the identification of conserved sequence motifs among the increasing number of known polymerase sequences. These efforts culminated in an alignment that included most DNA-directed DNA polymerases (DNAPs), reverse transcriptases (RTs), RNA-directed RNA polymerases (RNAPs) and DNA-directed RNAPs (Ref. 2). However, the tenuous nature of many of these sequence similarities cast doubt on the entire scheme, which therefore awaited confirmation and refinement, or rejection, based on further structure–function or structural studies. The past few years have seen the emergence of the structures of three new polymerases[3–9], which now make it possible to evaluate the significance of these patterns of apparent sequence conservation.

Motif conservation

The pattern of polymerase motif conservation identified by Delarue *et al.*[2], Poch *et al.*[10] and Mendez *et al.*[11] can be seen in Fig. 1. Although more extensive patterns of sequence similarity within certain polymerase families have been identified, we focus here on the limited set of motifs that are most widely distributed. It can be seen that there is a correlation between polymerase template or substrate specificity and motif conservation. For example, motifs T/DxxGR and B are found only in polymerases that use DNA templates, while motifs B′ and D are restricted to polymerases that use RNA templates and, within the RNA-directed family of polymerases, motif E is restricted to polymerases that use dNTPs. Motifs A and C unify the RNA- and DNA-directed RNA or DNA polymerases because they occur in polymerases of either template or substrate specificity.

Relating sequence motifs to structure

It is instructive to examine the overall structures of these enzymes before looking at where the sequence motifs occur within them (Fig. 2). The similarity in the shape of the polymerase domains of T7 RNAP, p66 HIV-1 RT, and DNAP I to a 'cupped right hand' has led to the designation of the three subdomains of the polymerase domain as 'fingers', 'palm', and 'thumb'[4]. The most extensive similarity is seen between T7 RNAP and DNAP I : the folding and almost all of the secondary structure in their respective polymerase domains is nearly identical, while the structural similarity with HIV-1 RT is limited to a core comprising most of the palm subdomain.

Motifs A and C. Peering more deeply into the large template-binding clefts of these enzymes, we can localize active sites that have been defined by structure–function studies, structural studies of polymerase–substrate/template complexes, and sequence comparison (Fig. 3). Most of the residues forming these active sites are part of the sequence motifs shown in Fig. 1. Motifs A and C form three strands of a β-sheet and a short segment of α-helix within the core of the palm subdomain, which is structurally similar in RT, T7 RNAP and DNAP I. Two amino acids (Asp537/Asp812 in T7 RNAP; Asp705/Asp882 in DNAP I; Asp110/Asp185 in HIV-1 RT), which are identified as invariant within these motifs, are brought into alignment when the three polymerases are superimposed. These two Asp residues bind and present two metal ions in the appropriate geometrical arrangement to catalyse a phosphoryl transfer reaction at the active site[12]. A third well-conserved carboxylate (Glu883 in DNAP I; Asp186 in RT; absent in T7 RNAP) is also expected to be involved in catalytic metal binding. Significantly, mutation of this third carboxylate reveals that it is less critical for activity than either of the two invariant Asp residues[13,14].

Motifs B and B′ are located in the fingers subdomains of DNAP I, T7 RNAP and RT, respectively. While motifs B and B′ are dissimilar in sequence and structure, and occur within a subdomain that is structurally dissimilar in the RNA-directed versus DNA-directed polymerases, they are similarly positioned relative to the center of the active site in both classes of polymerases. In the structure of HIV-1 RT complexed with primer–template, the fingers subdomain and elements from motif B′ contact the template strand[5]. Modeling, structural and mutation studies imply that a region in the corresponding position (including elements of motif B) of the fingers subdomain of DNAP I or T7 RNAP would be similarly involved in contacts with the template strand[15]. As the template in the RT primer–template structure does not extend downstream of the 3′ end of the primer, downstream template contacts must be deduced from modeling. The more compelling model would place the downstream template contacts on β-strands 3 and 4 of RT, the loop between these strands and (perhaps) the carboxy-terminal region

R. Sousa is at the Department of Biochemistry, Floyd Curl Drive, University of Texas Health Sciences Center, San Antonio, TX, USA.

© 1996, Elsevier Science Ltd

PII: S0968-0004(96)10023-2

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

of β-strand 11a (Ref. 16). However, it is also possible that these elements are involved in substrate contacts[17,18]. The latter hypothesis would be consistent with recent evidence from structural and mutational analyses that the amino-terminal region of motif B (helix O) in DNAP I interacts with the dNTP phosphates and ribose moiety[19,20]. The docked dNTP modeled by Arnold and colleagues in the RT primer–template complex would not contact the fingers subdomain, but would instead establish contact with elements of motifs A and C and (possibly) with β-strand 11 (Ref. 16). It is, therefore, unclear if the fingers subdomain is involved in substrate as well as template strand contacts, or if this represents a case where analogous structures in different polymerases have different functions (i.e. a role in substrate binding for the fingers of DNAP I and T7 RNAP, but not for RT).

Functional roles of the T/DxxGR motif and motif E. Irrespective of the question of substrate binding, it is clear that the fingers subdomains and elements of motifs B and B′ are involved in template-strand binding in both the DNA-directed and RNA-directed polymerases. It is, therefore, intriguing that structural similarity in the fingers subdomain and conservation of motifs B and B′ reflect polymerase template specificity. In the same way, we can examine the location and proposed function of the T/DxxGR motif, which occurs in the DNA-directed polymerases, but not in the RNA-directed polymerases. Mutational studies and modeling of template–DNAP I or template–T7 RNAP structures based on the RT primer–template complex reveal that the structure formed by this motif is also involved in template-strand contacts[15]. Along similar lines, it may be noted that motif E, the only one of five motifs conserved in RNA-directed DNA polymerases that is not also conserved in the RNA-directed RNA polymerases, forms a structure designated the 'primer grip', which is intimately associated with the primer strand[5].

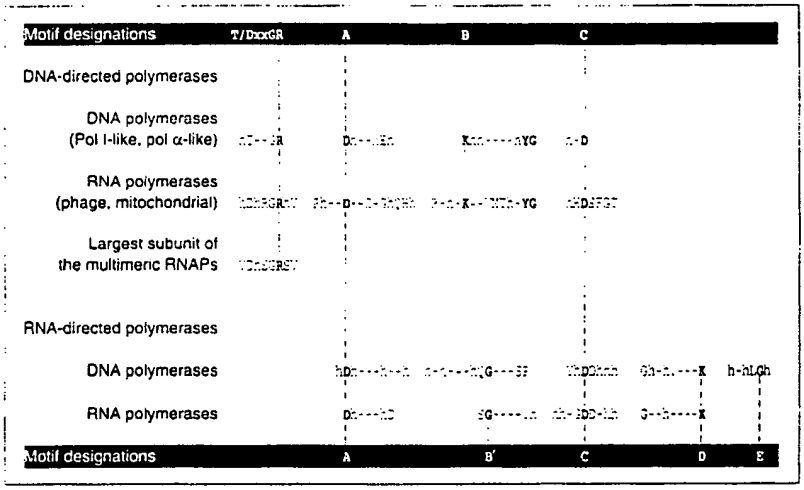


Figure 1

Patterns of motif conservation in nucleic acid polymerases[2,10,11]. Residues in blue are invariant. Other residues given are well conserved: h, hydrophobic residue; +, positively charged residue; –, any residue; .. a sequence gap. Residue numbers of invariant residues in DNA-directed DNA polymerase (DNAP) I, T7 DNA-directed RNA polymerase (RNAP) and HIV-1 reverse transcriptase (RT) are: for T/DxxGR motif – DNAP I, Arg668; T7 RNAP, Arg425; for Motif A – DNAP I, Asp705; T7 RNAP, Asp537; RT, Asp110; for Motif B – DNAP I, Lys758/Tyr766/Gly767; T7 RNAP, Lys631/Tyr639/Gly640; for Motif B′ – RT, Gly152; for Motif C – DNAP I, Asp882; T7 RNAP, Asp812; RT, Asp185; for Motif D – RT, Lys220; and for Motif E – RT, Gly231.

Structural differences between T7 RNAP and DNAP I. It is intriguing that there is one region where DNAP I and RT are more similar to each other than to T7 RNAP, even though DNAP I and T7 RNAP show greater structural similarity overall. RT and DNAP I both exhibit a fourth β-strand (β-strand 14 in DNAP I; 11a, 11b in RT), which extends the three-stranded sheet formed by motifs A and C. T7 RNAP lacks this fourth β-strand. As DNAP I and RT share substrate (and

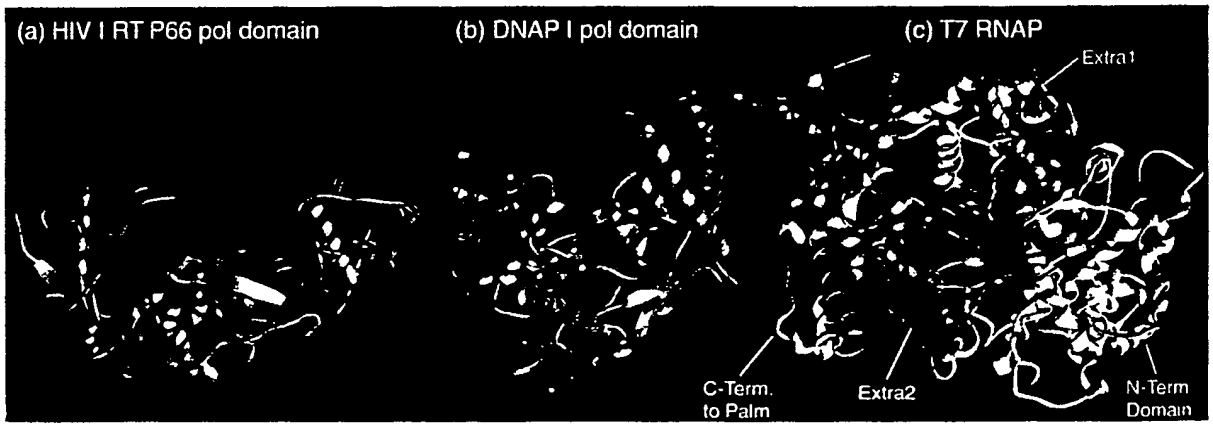


Figure 2

Structures of the polymerase domains of **(a)** P66 reverse transcriptase (RT), **(b)** DNA-directed DNA polymerase (DNAP) I and **(c)** the complete T7 DNA-directed RNA polymerase (RNAP) molecule[1,3–9]. The 'thumb' subdomains are colored green, the 'palm' subdomains are in red, and the 'fingers' subdomains are blue. Structural elements in T7 RNAP that have no counterpart in the DNAP I polymerase domain are colored light gray ('Extra1', 'N-Term' domain, 'C-Term to Palm') or orange ('Extra2'). The single magenta-colored helix in DNAP I and T7 RNAP is not formally considered part of the polymerase subdomain, but is conserved between T7 RNAP and DNAP I. The two green-colored spheres mark the positions of the invariant Asp residues, which identify the center of the active site.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

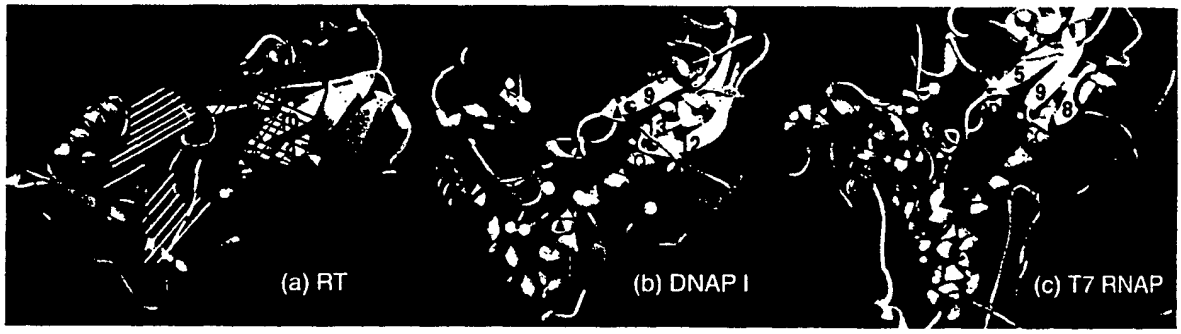


Figure 3

Polymerase domain structures of **(a)** reverse transcriptase (RT), **(b)** DNA-directed DNA polymerase (DNAP) I with the 'thumb' subdomains removed to allow an unobstructed view into the active site, and **(c)** the T7 DNA-directed RNA polymerase (RNAP) structure with the thumb and amino-terminal domains removed. The polymerases are presented with the long axis of the template-binding cleft parallel to the long axis of the page. RT helices E, F and strands 6, 9 and 10 correspond to DNAP I helices Q, R and strands 9, 12, 13, and to T7 RNAP helices CC, DD and strands 5, 8 and 9, respectively. The structures formed by motifs A and C are in yellow; motifs B and B′ are colored magenta; and the T/DxxGR motif (strands 3 and 4 in T7 RNAP; 7 and 8 in DNAP I) is orange. Motif E is colored green. Elements of the 'palm' not part of these motifs are in red, and the 'fingers' are blue. The carboxyl terminus of T7 RNAP is marked by an asterisk. T7 RNAP structures not a part of the polymerase domain are in light gray or orange. The region of RT in contact with the template strand in the RT–primer–template structure is indicated by diagonal white lines[5]. The diagonal yellow lines indicate where additional template-strand contacts might be made if the template strand were extended beyond the length in the crystal structure of the complex[16]. The region in contact with the primer strand is in indicated by diagonal blue lines[5]. Elements of the RT molecule that could contact a substrate dNTP in the RT active site are indicated by diagonal magenta lines[16]. Green triangles indicate the locations of mutations that affect NTP K_m* or nucleoside analog utilization, but whose effects are thought to be mediated by effects on template binding[21]. Magenta triangles indicate mutations that affect NTP K_m* or nucleoside analog utilization or become crosslinked to NTP, and whose effects are interpreted to be owing to direct NTP contacts[13–15,19–21,49,50]. Blue circles indicate the positions of mutations that affect K_d* for template or become crosslinked to the template and whose effects are interpreted to be owing to template-strand contacts[13–15,51]. Magenta circles indicate mutations that increase template K_d* and whose effects are interpreted to be owing to primer-strand contacts. The green spheres indicate the positions of the catalytic Asp residues. K_m* and K_d effects are fivefold or greater.

product) specificity, one possibility is that this structural divergence between the DNAPs and the RNAP is related to substrate/product specificity. Such a speculation is supported by the proposal that the AZT-resistance mutations at Lys219 of RT β-strand 11a exert their effects directly through contact with the dNTP (Ref. 21). It is also supported by the observation that mutations of Phe882 at the carboxyl terminus of T7 RNAP (which superimposes on Lys219 of RT) increase rNTP K_m (Ref. 22).

Alternatively, this structural pattern might be related to the fact that T7 RNAP uses a double-stranded template while DNAP I and RT use partially single-stranded templates. A fourth β-strand positioned in T7 RNAP analogously to the β-strand seen in DNAP I and RT could clash sterically with the domain we call 'Extra1', which is present in T7 RNAP, but not DNAP I or RT, and which might be involved in unwinding. This fourth β-strand could also occlude a groove in T7 RNAP in which the unwound non-template strand might bind. Truncation of the carboxyl terminus of T7 RNAP to remove the fourth β-strand might then reflect the need to remove structures that would sterically clash with the unwound non-template strand or with the protein domains responsible for unwinding.

Summary of motif structure–function correlations. Within a superfamily that includes the DNA pol I class of enzymes, the phage and mitochondrial RNAPs, the majority of RNA-directed polymerases, and (perhaps) the pol α class of enzymes, we can catalogue a set of correlations between conservation of structural elements and the functions of those elements. Motifs A and C, and most of the palm subdomain, are conserved irrespective of polymerase–template or substrate specificity, reflecting a direct role for these structures in the activity common to all polymerases: phosphodiester bond formation. Conservation of motifs T/DxxGR, B and the fingers domain in the DNA-directed polymerases, and conservation of motif B′ and a distinct fingers domain in the RNA-directed polymerases reflects a role for these elements in template-strand binding. The unique position of Motif E, as the motif that is conserved only in the DNA-synthesizing polymerases within the RNA-directed class of enzymes, reflects the role of this structural element in product (i.e. primer) contacts. Finally, structural similarity in β-strands 14 and 11 of DNAP I and RT, respectively, and divergence with T7 RNAP in the corresponding region might reflect a role for this structural element in substrate/product contacts or the structural requirements for utilization of a double-stranded template.

Modularity in polymerase architecture

One way to look at polymerase structure is to imagine that polymerases are assembled from modules whose structural conservation often reflects common function, as illustrated in Fig. 4. The fingers, palm and thumb subdomains together form the polymerase domain. The thumb subdomains, whose functions have not yet been addressed, are the least well conserved of these three subdomains, but appear to be similar in certain general features: they are extended, flexible, predominately α-helical structures that are involved in conferring processivity on polymerization by wrapping around bound template and/or interacting directly with the template to inhibit polymerase–template dissociation[23–26].

The function of the complete polymerase domain appears limited to the minimal function of template-directed processive polymerization. Specific polymerases can display additional activities, but these activities appear to reside on distinct domains. Thus, T7 RNAP is capable of sequence-specific (promoter) DNA binding and template unwinding. It emerges that promoter-specific binding may be largely conferred on T7 RNAP

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

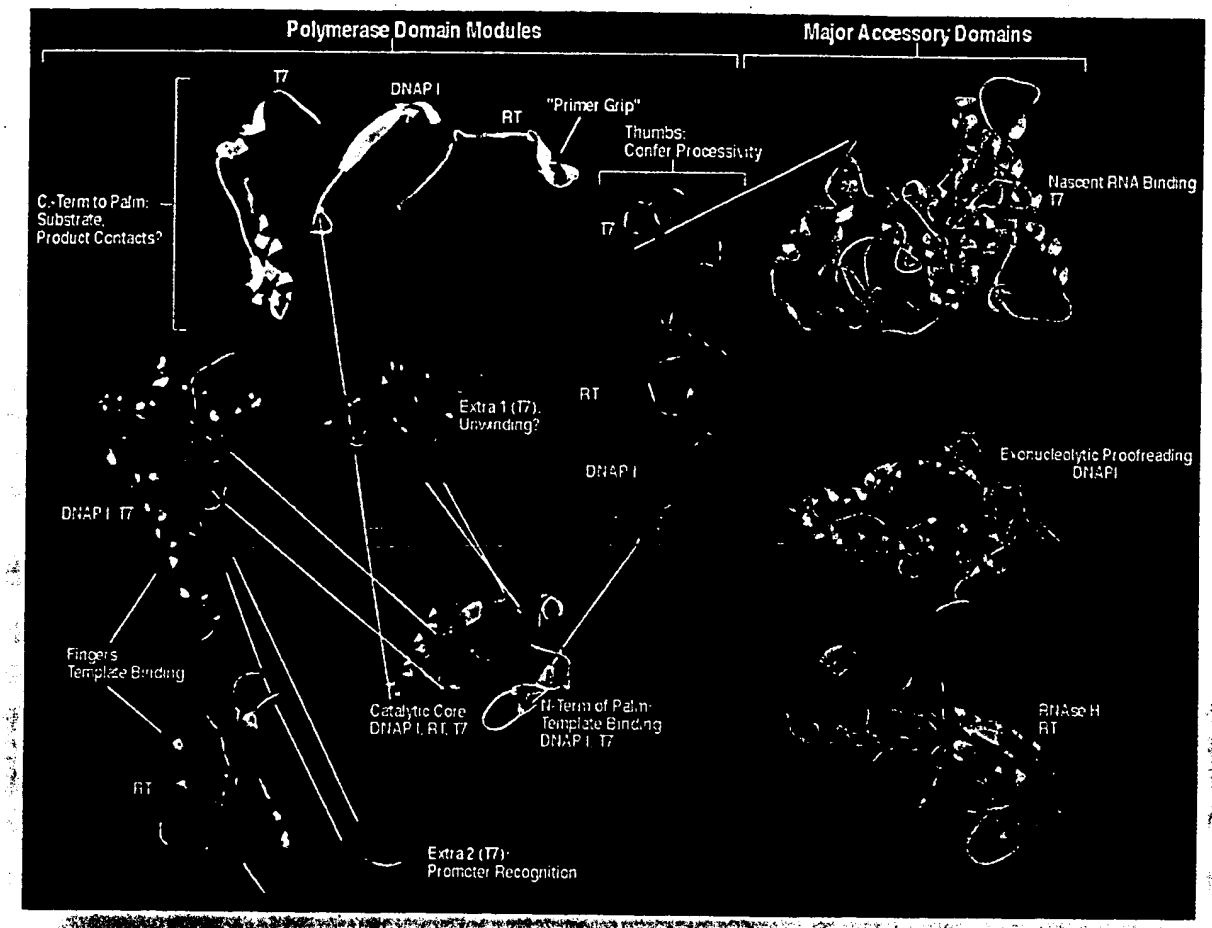


Figure 4

A modular architecture for polymerases is suggested. The white lines indicate how different structural elements are linked to each other. The functions of these structural elements and their occurrence in different polymerases are indicated.

by the 'Extra2' domain (Fig. 4), as amino acids critical for promoter recognition map to this loop[27,28]. Template-unwinding activity in T7 RNAP has not been mapped, but it is possible that the Extra2 domain plays a role in this process as it is placed at the leading end of the polymerase and forms two grooves into which the unwound strands of DNA might fit. The major accessory domains of each of these polymerases (RNAse H in RT; Exo proofreading in DNAP I; amino-terminal domain in T7 RNAP) are responsible for distinctive activities in each enzyme and bear no structural similarity to one another, although they do occupy roughly equivalent positions so that they can interact with nucleic acid strands upstream of the polymerase active site.

The scheme for polymerase organization presented in Fig. 4 suggests that gene fusion and recombination events might have played a role in the evolution of modern multidomain polymerases from multi-subunit enzymes composed of more simple polypeptides. In this light, the structure of the middle RNAP of phage N4 is intriguing because it shows homology to the complete T7 RNAP polymerase domain, but is composed of two subunits (designated P4 and P7; L. Rothman-Denes, pers. commun.). The break in the N4 RNAP places the amino terminal of the palm and thumb subdomains onto one subunit, and the conserved catalytic core and fingers structures on the other. Thus, in this instance, the subunit separation coincides with one of the subdomain divisions presented in Fig. 4.

Structure and extent of this polymerase superfamily

Using the alignments of Delarue *et al.*[2] and Poch *et al.*[10] as a guide, it is expected that the pol I class of DNAPs and the phage and mitochondrial RNAPs will display polymerase domain structures with palms and fingers similar to those shown in Fig. 4, while the structures of the thumb subdomains might show variation within the range revealed in this figure. The alignment by Delarue *et al.* included the pol β class of DNA polymerases in its scheme for unification, but also suggested that the pol β class was exceptionally divergent because it was unique among the DNA-directed polymerases in lacking an identifiable motif B sequence. Analysis of the X-ray structure of pol β suggests that, in fact, this polymerase is not related to DNAP I, RT or T7 RNAP and might be more closely related to nucleotidyl transferases, which are not polymerases[29]. The pol α class of DNA polymerases is expected to display a palm like that of DNAP I or T7 RNAP, however, changes in the spacing between motifs A, B and C imply at least some topological rearrangements in the fingers domain of the pol α class of polymerases relative to the pol I class. For the RNA-directed polymerases,

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

analysis of the spacing between motifs also suggests that the palm domain structures will be relatively invariant, while there might be variations in the fingers-domain structures of the most distantly related classes. Alternatively, variations in motif spacing could be accommodated in 'extra' modules and variation in lengths of different elements on the periphery of these domains, which would leave the folding of their cores unchanged (as is seen, for example, in the comparison of the T7 RNAP and DNAP I structures).

Status of the multi-subunit RNAPs

Not included in the Delarue *et al.* alignment were the multi-subunit RNAPs (prokaryotic RNAPs, eukaryotic RNAPs I, II and III). The identified sequence similarities between the catalytic subunits of these RNAPs and polymerases of the RT–DNAP I–T7 RNAP superfamily[11,30,31] are not extensive enough to be conclusive without confirmation from structures at atomic resolution and, despite observations of similarity in the overall shape of the multi-subunit RNAPs and T7 RNAP[32–34], it might be more profitable to compare the extensive mechanistic, rather than the structural, similarities of these enzymes. These include similarities in the timing, triggering and conformational changes involved in the transition from the poorly processive initiation phase of transcription to the elongation phase[35–41], and the utilization of similar promoter[42,43] and terminator sequences[44–47] by both classes of polymerases. The observation that the yeast RNAP holoenzyme is a dimer composed of a core enzyme that is homologous to T7 RNAP and a polypeptide that is functionally and structurally similar to the sigma subunit of *Escherichia coli* RNAP[48], might also be relevant to understanding the relationship between the single subunit and multi-subunit RNAPs.

Concluding remarks

The significance of the observed pattern of structural similarities between nucleic acid polymerases, the correlations between structural similarity and the functional role of different structural elements, and the mechanistic similarities between the multi-subunit RNAPs and simpler polymerases like T7 RNAP have yet to be fully explored. Do the observed patterns reflect strict functional requirements for the utilization of different templates or substrates, or do they reflect contingent events in the evolutionary history of polymerases? To what degree can the different functionalities displayed by a polymerase like T7 RNAP be distinctly mapped to different structural modules? Studies aimed at answering these questions and others might profit from a perspective that integrates the recognized structure–function relationships of nucleic acid polymerases.

Acknowledgements

Work in the author's laboratory is supported by NIH grant GM52522-01A1.

References

1 Ollis, D. L., Brick, P. and Steitz, T. A. (1985) *Nature* 313, 765–769
2 Delarue, M. *et al.* (1990) *Protein Eng.* 10, 461–467
3 Arnold, E. *et al.* (1992) *Nature* 357, 85–89
4 Kohlstaedt, L. A. *et al.* (1992) *Science* 256, 1783–1790
5 Jacobo-Molina, A. *et al.* (1993) *Proc. Natl Acad. Sci. USA* 90, 6320–6324
6 Sousa, R., Chung, Y. J., Rose, J. R. and Wang, B. C. (1993) *Nature* 364, 593–599
7 Pelletier, H. *et al.* (1994) *Science* 264, 1891
8 Sawaya, M. R. *et al.* (1994) *Science* 264, 1930
9 Ren, J. *et al.* (1995) *Nat. Struct. Biol.* 2, 293–302
10 Poch, O., Sauvaget, I., Delarue, M. and Tordo, N. (1989) *EMBO J.* 8, 3867–3874
11 Mendez, L., Blanco, L., Lazaro, J. M. and Salas, M. (1994) *J. Biol. Chem.* 269, 30030–30038
12 Steitz, T. A., Smerdon, S. J., Jager, J. and Joyce, C. M. (1994) *Science* 266, 2022–2025
13 Polesky, A. H., Steitz, T. A., Grindley, N. D. and Joyce, C. M. (1989) *J. Biol. Chem.* 265, 14579–14591
14 Polesky, A. H., Steitz, T. A., Grindley, N. D. F. and Joyce, C. M. (1992) *J. Biol. Chem.* 267, 8417–8428
15 Joyce, C. M. and Steitz, T. A. (1994) *Annu. Rev. Biochem.* 63, 777–822
16 Patel, P. H. *et al.* (1995) *Biochemistry* 34, 5351–5363
17 Cheng, N. *et al.* (1993) *Biochemistry* 32, 7630–7634
18 Wu, J. *et al.* (1993) *J. Biol. Chem.* 268, 9980–9985
19 Astatke, M., Grindley, N. D. F. and Joyce, C. M. (1995) *J. Biol. Chem.* 270, 1945–1954
20 Beese, L. S., Friedman, J. M. and Steitz, T. A. (1993) *Biochemistry* 32, 14095–14101
21 Tantillo, C. *et al.* (1994) *J. Mol. Biol.* 243, 369–387
22 Patra, D., Sousa, R. and Lafer, E. M. (1992) *J. Mol. Biol.* 224, 307–318
23 Bonner, G., Lafer, E. M. and Sousa, R. (1994) *J. Biol. Chem.* 42, 25120–25128
24 Sousa, R., Rose, J. and Wang, B. C. (1994) *J. Mol. Biol.* 225, 239–252
25 Beard, W. A. *et al.* (1994) *J. Biol. Chem.* 269, 28091–28097
26 Beese, L. S., Derbyshire, V. and Steitz, T. A. (1993) *Science* 260, 352–355
27 Joho, K. E. *et al.* (1990) *J. Mol. Biol.* 215, 21–29
28 Raskin, C. A., Diaz, G., Joho, K. and McAllister, W. T. (1992) *J. Mol. Biol.* 228, 506–516
29 Holm, L. and Sander, C. (1995) *Trends Biochem. Sci.* 20, 345–347
30 Hermann, T., Meier, T., Gotte, M. and Heumann, H. (1994) *Nucleic Acids Res.* 22, 4625–4633
31 Kim, W. J., Burke, L. P. and Martin, M. A. (1994) *J. Mol. Biol.* 244, 13–22
32 Darst, S. A., Kubalek, E. W. and Kornberg, R. D. (1989) *Nature* 340, 730–732
33 Darst, S. A., Edwards, S. M., Kubalek, E. W. and Kornberg, R. D. (1991) *Cell* 66, 121–135
34 Schultz, P. *et al.* (1993) *EMBO J.* 12, 2601–2607
35 Carpousis, A. J., Gralla, J. D. (1980) *Biochemistry* 19, 3245–3253
36 Martin, C. T., Muller, D. K. and Coleman, J. E. (1988) *Biochemistry* 27, 3966–3974
37 Krummel, B. and Chamberlin, M. J. (1989) *Biochemistry* 28, 7829–7842
38 Ikeda, R. A. and Richardson, C. C. (1987) *J. Biol. Chem.* 262, 2800–3808
39 Hansen, U. M. and McClure, W. R. (1980) *J. Biol. Chem.* 255, 9564–9570
40 Sousa, R., Patra, D. and Lafer, E. M. (1992) *J. Mol. Biol.*, 224, 319–334
41 Polyakov, A., Severinova, E. and Darst, S. A. (1995) *Cell* 83, 365–373
42 Jaehning, J. A. (1993) *Mol. Microbiol.* 8, 1–4
43 Sandig, V., Lieber, A., Bahring, S. and Strauss, M. (1993) *Gene* 131, 255–259
44 Farnham, P. J. and Platt, T. (1980) *Cell* 20, 739–748
45 Jeng, S-T., Gardner, J. F. and Gumport, R. I. (1990) *J. Biol. Chem.* 265, 3823–3830
46 Reynold, R. and Chamberlin, M. J. (1992) *J. Mol. Biol.* 224, 53–63
47 Jeng, S-T., Gardner, J. F. and Gumport, R. I. (1992) *J. Biol. Chem.* 267, 19306–19312
48 Mangus, D. A., Jang, S. H. and Jaehning, J. A. (1994) *J. Biol. Chem.* 269, 26568–26574
49 Sousa, R. and Padilla, R. (1995) *EMBO J.* 14, 4609–4621
50 Osumi-Davis, P. A. *et al.* (1994) *J. Mol. Biol.* 237, 5–19
51 Rechinsky, V. O. *et al.* (1995) *Mol. Gen. Genet.* 247, 110–113

TiBS newsgroup

Did you know that there are hundreds of visitors every month to the *TiBS* newsgroup?

Make your views heard!

The *TiBS* newsgroup would welcome comments, questions and discussion on articles and topics covered in *TiBS*, teaching tips and problems, nomenclature issues, the impact of the Internet on science and scientific publishing, or on any topic relevant to biochemistry and molecular biology.

To send a contribution to the newsgroup, Email your message to:

tibs@daresbury.ac.uk
(Europe, Africa or central Asia)

or

tibs@net.bio.net
(Americas or the Pacific Rim)

Please specify the theme of your message clearly in the 'subject' heading and include a brief introduction in the body of your posting (for example, mention the *TiBS* article to which you are referring). This will allow the discussion to be followed more easily on the newsgroup.

This Document has been provided by Technical Information Service of Purdue University and is copyright protected.

DNA Polymerase Structure and Mechanisms of Action

Manju M. Hingorani[1] and Mike O'Donnell*[1,2]

[1]*The Rockefeller University and* [2]*The Howard Hughes Medical Institute, New York, New York 10021, USA*

*Address correspondence to this author at the The Rockefeller University, and The Howard Hughes Medical Institute, New York, New York 10021, USA; Phone: (212) 327 7251; Fax: (212) 327 7254; email: odonnel@rockvax.rockefeller.edu

Abstract: The ubiquitous and essential nature of DNA predicts that enzymes responsible for DNA synthesis evolved early and share a common design and mechanism of action. DNA polymerases from many different organisms do exhibit striking similarities in their overall architecture, the design of the catalytic site, and the mechanism of nucleotidyl transfer. In spite of these shared features, however, DNA polymerases display an astonishing variety in structure and function, ranging from single subunit enzymes specialized for DNA repair to complex multi-protein holoenzymes responsible for genomic DNA replication. This review uses a few well-characterized model systems to summarize our current understanding of the workings of DNA polymerases in DNA metabolism.

INTRODUCTION

DNA polymerases catalyze the synthesis of deoxyribonucleic acid in a template-dependent process that results in a faithful copy of the original DNA molecule [1]. Consequently, these enzymes are necessary to propagate, maintain, and manipulate the genetic code of living organisms. The discovery of Polymerase I (Pol I) from *E. coli* in 1956 [2,3], and the crystal structure of the Pol I Klenow fragment in 1985 [4], heralded an explosion in DNA polymerase research, which continually yields new and surprising insights into the structure and mechanisms of these enzymes. The basic chemistry of nucleotidyl transfer for extending the DNA polymer one nucleotide at a time is universal, and all polymerases appear to utilize a similar catalytic mechanism for DNA synthesis [5]. However, an extreme diversity of DNA polymerases exists, not just among different organisms but within individual organisms. In *Saccharomyces cerevisiae*, for example, 8 different DNA polymerases had been discovered at the last count [6]. These enzymes function alone or with accessory factors, and vary substantially in the fidelity, speed, and processivity with which they replicate DNA. DNA polymerases are necessary not only for genomic DNA replication, but also for repair of DNA damage, DNA recombination, as well as the link between these processes and cell cycle regulation. Thus, the diversity in polymerases likely reflects the wide variety of metabolic situations in which DNA synthesis is required.

Based on their functions, DNA polymerases can be broadly classified into two groups: A) replicative DNA polymerases (DNA replicases), that are responsible primarily for duplicating genomic DNA and B) repair polymerases, that primarily fix damaged DNA (including the new

family of polymerases responsible for DNA lesion bypass). Replicative DNA polymerases must synthesize extended lengths of DNA with high speed as well as high accuracy to ensure that each daughter cell receives a true copy of the genetic code on cell division. In general, these DNA replicases are complex macromolecular assemblies of several proteins that function together for efficient DNA replication [7,8]. In contrast to DNA replicases, polymerases dedicated to DNA repair generally have a simpler architecture and appear designed for DNA synthesis localized to areas of DNA damage. For example, replicative DNA polymerases tend to stall or dissociate from the template when they encounter lesions caused by ultraviolet radiation or chemical damage to DNA. On these occasions, specialized repair polymerases access the site and synthesize short DNA fragments to bypass the lesion, following which DNA replicases restart replication [9,10]. The DNA lesions are ultimately removed by DNA repair processes such as nucleotide excision repair [11], base excision repair [12-15], and mismatch repair [16,17] that employ several proteins including both repair-dedicated polymerases (e.g., *E. coli* Pol I in nucleotide excision repair) and replicative DNA polymerases (e.g., *E. coli* Pol III in mismatch repair).

A more structure-based classification of DNA polymerases groups them into several different families based on similarities in amino acid sequence and structure [18,19]. The DNA polymerase I or Pol A family includes bacterial and bacteriophage polymerases such as *E. coli* Pol I, *Thermus aquaticus* DNA polymerase, and the bacteriophage T7 DNA polymerase, among others. These enzymes function in DNA replication (e.g., T7 polymerase, [20]), and repair (e.g., Pol I in nucleotide excision repair [11]). The polymerase B or Pol α family is more diverse and comprises bacterial enzymes (e.g., *E. coli* Pol II [1]), eukaryotic DNA replicative polymerases such as Pol α, Pol δ, and Pol ε [6], replicative polymerases from bacteriophage T4 and RB69 [21], as well as the archaeal Pol B polymerases, among others [22,23]. The *E. coli* DNA polymerase III (α subunit) shares little sequence homology with any other polymerase and is therefore in a class by itself (polymerase C family) [24]. Polymerase β (family X), also unrelated to the above polymerases, is an eukaryotic enzyme responsible for filling in single nucleotide gap in DNA during base excision repair [12,13]. Recently identified archaeal polymerases related to Pol β are also included in family X [22,23]. The RT family contains RNA-dependent DNA polymerases including HIV reverse transcriptase and telomerases [25]. A new, rapidly growing family of polymerases contains DNA polymerases such as the *E. coli* UmuC [26] and DinB proteins [27], and the *S. cerevisiae* Pol η and Pol ζ [10], that are responsible for lesion bypass [28,29]. Finally, the DP2 family (named after the *Pyrococcus furiosus* DP2 polymerase [30]) contains newly identified archaeal DNA polymerases that share little homology with the enzymes listed above [22,23]. These families are subdivided further according to particular similarities and differences in polymerase structure and function, highlighting the incredible diversity of these enzymes. In this review we focus on a few well-characterized polymerases to summarize our current understanding of the structure and mechanism of action of the catalytic polymerase protein as well as DNA polymerase holoenzymes (in which several proteins

function together to replicate genomic DNA).

DNA POLYMERASE STRUCTURE AND FUNCTION

Polymerases are unusual enzyme catalysts in that they use DNA, a substrate in the reaction, as a template to direct formation of the product, which is also DNA. In order to make a true copy of DNA, these enzymes are designed to: 1) recognize and bind a primed DNA strand, which serves as the template for synthesis, 2) recognize and bind a deoxynucleoside 5'-triphosphate (dNTP) and match it with the template to form a base pair, 3) catalyze nucleotidyl transfer in which a covalent link is formed between the 3'-hydroxyl group on the primer and the incoming dNTP, and finally 4) reposition the newly extended DNA polymer for the next cycle of nucleotidyl transfer [1]. DNA polymerases can also proofread the newly synthesized DNA with a 3' – 5' exonuclease activity, which reduces their inherently low error rate of 1 in 10^5 base pairs to that approaching 1 in 10^8 base pairs during DNA replication [31-35].

Fig. (1). The shape of a DNA polymerase. The polymerase domain of the *E. coli* polymerase I Klenow fragment represents the typical topology of a DNA polymerase. The fingers, palm, and thumb subdomains are arranged to form a deep DNA binding cleft, with the catalytic site in the palm, at the base of the cleft. The proofreading exonuclease lies adjacent to the thumb domain and extends south of the polymerase domain in this view. The position of the exonuclease domain varies with respect to the polymerase domain in different enzymes (Pol I [4] and T7 polymerase [37] *versus* RB69 [71] and D. tok polymerases [40]).

The DNA replication reaction presents several challenges for an enzyme catalyst. Predominant among these are the requirement that polymerase bind all four dNTPs and yet distinguish the correct dNTP from the other three in each catalytic cycle. The polymerase also has to bind DNA with high affinity and release it—but without complete dissociation—at the end of each reaction, so that it can be repositioned within the active site for the next reaction. Remarkably, DNA polymerases are capable of performing this complex task with high speed and fidelity. In the past four decades, the structure and kinetic mechanisms of DNA polymerases have been the subject of intense study, in an effort to understand how these biological machines work. The first crystal structure of a DNA polymerase (Klenow fragment of *E. coli* Pol I) yielded valuable insights into how polymerases are designed to catalyze DNA replication (Fig. 1) [4,19]. Subsequently, crystal structures of several other polymerases revealed that despite differences in amino acid sequence and structure among enzymes from different families, they share several features that are essential for activity (e.g., Pol I family: Klenow fragment [4], *Thermus aquaticus* pol [36], T7 DNA pol [37]; *Bacillus stearothermophilus* pol [38]; Pol II family: RB69 pol [39], *D. tok.* Pol [40]; family X: pol β [41]; RT family: HIV-RT [42]. The shape of the polymerase domain is by far the most prominent feature common to all the polymerase structures determined to date. As described first for the Klenow fragment, the polymerase resembles a half-open right hand with the "palm" subdomain forming a cleft that is flanked by the

"fingers" and "thumb" subdomains. Together, the three subdomains hold the primer-template DNA and position the incoming dNTP for incorporation into DNA. The palm subdomain contains the catalytic site where the chemistry of nucleotidyl transfer takes place. The fingers subdomain interacts with and positions the template DNA strand and the incoming dNTP, while the thumb subdomain primarily binds the duplex DNA in a sequence-independent manner along the minor groove. Most DNA polymerases also contain a 3' - 5' exonuclease domain within the same polypeptide, which proofreads newly synthesized DNA and corrects mismatched base pairs [5,43,44].

DNA Polymerase Mechanism

Initial research on DNA polymerase mechanisms was also focused on the Klenow fragment (an enzyme specialized for DNA repair), and the use of rapid kinetic techniques revealed individual steps in the polymerase reaction pathway [45-48]. Further analyses of such enzymes as the T7 DNA polymerase [34], HIV reverse transcriptase [49], and the T4 polymerase [50], yielded the detailed mechanism by which replicative DNA polymerases synthesize DNA. Fig. **2** shows the polymerase reaction pathway with kinetic and thermodynamic parameters that define the catalytic activity of the T7 DNA polymerase. The core elements of this reaction mechanism generalize to the other DNA polymerases, although the equilibrium constants and rates of reaction vary [34].

Fig. (2). Kinetic mechanism of DNA polymerase activity. The pathway of DNA synthesis catalyzed by DNA polymerases, with kinetic and thermodynamic parameters defining the bacteriophage T7 polymerase mechanism. Both DNA and the correct dNTP bind polymerase rapidly and with high affinity, following which the polymerase undergoes a slow, rate-limiting conformational change (Pol_DNA_dNTP to Pol*_DNA _dNTP) in preparation for nucleotidyl transfer. The chemistry step is fast, and is followed by another conformational change in the polymerase, which leads to product release and translocation of DNA for the next catalytic cycle. Figure adapted from [51].

The polymerase first binds a primed DNA template with high affinity ($K_d \sim 20$ nM), and then binds a dNTP molecule to form a ternary Pol_DNA_dNTP complex [51]. The dissociation constant for interaction between polymerase and the correct nucleotide is 20 μM *versus* 4 to 8 mM for nucleotides that do not match the DNA template [51,52]. Such discrimination against incorrect nucleotides suggests that the initial dNTP-binding step involves base-pair formation between the nucleotide and template DNA, during which the polymerase can accept a good match or reject a mismatch. Similar selectivity in dNTP binding has been observed in the case of HIV reverse transcriptase [49] and the T4 DNA polymerase [53]. Next, there is a slow, rate-limiting step in the reaction, prior to the chemical step of nucleotidyl transfer, which results in formation of the **Pol*_DNA_dNTP** complex. The rate of dissociation of DNA from this complex (and an analogous HIV reverse transcriptase complex) is 10-fold slower than from its precursor, the Pol_DNA_dNTP complex (Fig. **2**) [34,49]. These data predicted that the polymerase changes conformation from an "open" state to a "closed"

state during the rate-limiting step, and clamps down on the DNA and dNTP in preparation for chemistry. Further investigation revealed that the enzyme goes forward with this conformational change only when a correct dNTP (which forms a Watson-Crick base pair with the template) is bound to the polymerase [51,52]. In the T7 polymerase, for example, at saturating concentrations of dNTPs, the incorporation rate of a correct dNTP is higher than the misincorporation rate by a factor of 2000 – 4000 [34]. The chemistry step remains hostage to the rate-limiting conformational change, which in turn is sensitive to the "correctness" of the new base pair formed by the nucleotide and template, and is slowed down by an incorrect nucleotide in the active site. In addition, the chemistry of nucleotidyl transfer is inhibited, contributing to the slow rate of misincorporation [52]. A partial proteolytic mapping study of the T7 DNA polymerase has also shown that the polymerase changes conformation only when bound to the correct nucleotide [54]. Thus, DNA polymerase uses an "induced-fit" mechanism for fidelity of nucleotide incorporation, in which a slow step in the reaction (predicted as a change in polymerase from "open" to "closed" state) severely restricts nucleotidyl transfer in the presence of an incorrect dNTP and a mismatched base pair at the active site. In contrast, after the correct nucleotide is bound and the polymerase changes conformation, the chemistry step occurs rapidly and extends the primer by one nucleotide. The polymerase then goes through a second conformational change that likely returns it to an "open" state, leading to pyrophosphate release and translocation of DNA for the next cycle of dNTP incorporation (Fig. **2**). The following sections describe particular structural features of DNA polymerase that play a critical role in the mechanism of DNA polymerization.

The Structural Basis for Polymerase Action

A) Catalysis

As described earlier, the nucleotidyl transfer reaction in which the 3' - hydroxyl group of the primer attacks the α-phosphate of the incoming dNTP occurs in the catalytic site located in the palm domain of the DNA polymerase. The active site, shown in Fig. **3A** (for T7 DNA polymerase [37]), contains several acidic and polar amino acid residues as well as two metal cations (usually Mg^{2+}) that are essential for the proposed mechanism of catalysis [19]. In particular, two aspartate residues are absolutely conserved between the polymerase families, and these provide the carboxylate oxygens that ligate the metal ions [39,43]. Ion A is located near the 3'-hydroxyl group of the DNA primer and the α-phosphate of the incoming dNTP. In this location, ion A is ideally positioned to lower the pK_a of the hydroxyl group and facilitate formation of the hydroxide anion which can initiate nucleophilic attack on the α-phosphate of dNTP (Fig. **3B**). Metal ion A also ligates the α-phosphate of dNTP, helping to stabilize the pentacovalent transition-state formed during the reaction. Metal ion B ligates oxygens in all three phosphate groups of the dNTP, likely helping align the triphosphate moiety for attack by the 3'-hydroxyl, as well as stabilizing the charge on the transition state. Other polar residues in the active site and possibly ion B help stabilize the charged pyrophosphate group as it dissociates

from the polymerase after nucleotidyl transfer is complete. This two-metal ion mechanism for nucleotidyl transfer was originally proposed by Steitz, based on an analogous mechanism determined for the 3' – 5' exonuclease [55]. Remarkably, the catalytic mechanism and the geometry of the active site are highly conserved even in completely unrelated polymerases such as the T7 DNA polymerase [37] and Pol β [56], reflecting the evolutionary success of this mechanism for DNA synthesis.

Fig. (3). The catalytic site. (A) The T7 DNA polymerase active site shows two metal ions (A and B) coordinated by conserved aspartate residues, a main chain oxygen, and water molecules, that in turn ligate the oxygens in the three phosphate groups of the incoming dNTP. (B) The 3' - hydroxyl group on the primer initiates nucleophilic attack on the α phosphate of the incoming dNTP, which results in phosphodiester bond formation between the primer and nucleotide and pyrophosphate release. Figure adapted from [37].

B) Selection of dNTP

In addition to the layout of the core catalytic site, the crystal structures reveal several important features that allow DNA polymerases to synthesize DNA with high fidelity and processivity. The ability of a DNA polymerase to faithfully complement a DNA template depends on how well it selects for correct pairing between the template and the nucleotide during initial base pair formation, during extension of the base pair, and finally on the editing activity of the 3' – 5' exonuclease. The crystal structures of polymerases in complex with DNA and an incoming nucleotide reveal how a polymerase might initially select the correct nucleotide (e.g., T7 DNA polymerase [37], pol β [41], HIV-RT [57], and *B. stearothermophilus* polymerase [38], among others). As the template DNA strand enters the active site, it bends away from the cleft between the fingers and thumb and wraps around the fingers subdomain. This conformation exposes the next base in the template for pairing with the incoming dNTP (Fig. **4** for the T7 polymerase). The dNTP-binding site is located in a narrow junction between the fingers and thumb domains (Fig. **4** for T7 polymerase and Fig. **5D** for Klentaq1 polymerase). The 3' - end of the primer lies right next to the dNTP-binding site, and together with residues from the fingers domain it forms a highly constrained binding pocket for the new base pair [37]. The T7 polymerase contacts the bases through numerous van der Waals' interactions and can therefore precisely recognize the geometry of the base pair. The DNA Pol β structure shows extensive hydrogen bonding between the polymerase active site and the new base pair, which also facilitates recognition [41]. Accordingly, DNA polymerases can select for correct, Watson-Crick base pairs and reject distorted, mismatched base pairs when the incoming dNTP initially fits into the binding pocket [58,59].

Fig. (4). T7 DNA polymerase•DNA•dNTP complex (with thioredoxin). The structure of the T7 DNA polymerase shows the primer-template in the DNA-binding cleft, with the primer 3' - end held in the polymerase active site, and the template bent to expose the next base for pairing with the incoming dNTP. The fingers domain presses close to the active site to form a tight binding pocket for the nucleotide and the new base pair. The thumb domain contacts the duplex region of the primer-template in the minor groove and increases the polymerase's

grip on DNA. The thioredoxin processivity factor appears twisted away from the DNA, and is perhaps in an open conformation in this structure. Figure adapted from [37].

Biochemical studies of DNA polymerase mechanism describe an intramolecular rate-limiting step in the reaction, after initial dNTP binding to the polymerase-DNA complex and before its incorporation into DNA (Fig. **2**) [34]. This step is predicted to be a conformational change that is induced by binding of the correct dNTP to the active site and results in a tightly bound **Pol***_DNA_dNTP complex that is competent for catalysis. The first clues regarding the structural nature of this step were provided by crystal structures of the Pol β_DNA_dNTP ternary complex, the Pol β_DNA binary complex, and the apo form of Pol β (free of ligands) [41,56,60]. Compared to the latter two structures, Pol β complexed with nucleotide and DNA has a relatively "closed" conformation in which domain D (the fingers domain) lies very close to the newly formed base pair and appears sensitive to the geometry of its fit in the active site. The T7 polymerase structure in complex with a nucleotide and primed DNA also has a relatively closed conformation when compared with related Pol I family polymerases that are either unliganded or bound to DNA only [37,61]. In the T7 Pol_DNA_dNTP complex, the fingers domain (most prominently the O helix) is angled in toward the palm domain and DNA (Fig. **4**), whereas in the other enzymes (apo form or Pol_DNA complexes) the fingers are angled away from the active site (~ 40_ difference between the two orientations), resulting in a more open conformation. The authors proposed that transition between these open and closed forms of the polymerase corresponds to the rate-limiting conformational change that DNA polymerase undergoes after binding the correct dNTP (Fig. **2**) [37,61].

Most recently, a series of structures of a single polymerase, the Klenow fragment of *Thermus aquaticus* DNA polymerase (Klentaq1) provided direct evidence for closure of the fingers domain toward the active site [62,63]. Fig. **5A** and **5B** show the structures of a Klentaq1_DNA binary complex and a **Klentaq1***_DNA_dNTP
ternary complex, respectively. The **Klentaq1***_DNA_dNTP structure was obtained from crystals that were analyzed immediately after growth in the presence of ddCTP and primed DNA, and thus represents a closed conformation in which the enzyme is associated tightly with both the DNA and the correct nucleotide (Fig. **5B**). The Klentaq1_DNA structure was obtained by depriving the closed ternary complex crystals of DNA and ddCTP for a month, and thus represents a conformation in which the enzyme is bound only to DNA, having released the nucleotide (Fig. **5A**). As expected from the earlier analysis of Pol β and Pol I family polymerase structures, the most striking difference between the two forms of Klentaq1 is in the conformation of the fingers domain. The Klentaq1_DNA form is fully open, with the tip of the fingers domain (O helix) twisted away from the body of the enzyme such that a crevice is visible at the top of the cleft formed by the fingers, thumb, and palm domains (Fig. **5A**). The **Klentaq1***_DNA_dNTP form is fully closed, such that the crevice is no

longer visible (Fig. **5B**). A third structure of Klentaq1 was obtained by depriving ternary complex crystals of DNA and ddCTP for a few days only. This procedure yielded an intermediate ternary complex (Klentaq1_DNA_dNTP) that contains nucleotide in the binding site, but has an open conformation similar to that of the Klentaq1_DNA binary complex. Superimposition of the two ternary complexes (Fig. **5C**) highlights the inward movement of the fingers domain as the polymerase changes from open to closed conformation. In the closed conformation, the O helix lies much closer to the active site and helps orient the DNA template for base-pairing with the incoming dNTP (Fig. **5D**). Furthermore, in the closed conformation the nucleotide is buried deep in the active site (in an environment favorable for chemistry), whereas in the open conformation, the nucleotide is readily accessible to the solvent (Fig. **5E**).

Fig. (5). Klentaq1 polymerase conformations. (A) An image of the open Klentaq1•DNA binary complex, with primed DNA bound in the cleft. The blue arrow predicts the motion of DNA during translocation. (B) The ternary Klentaq1*•DNA•dNTP complex shows the inward movement of the fingers domain, which closes the crevice at the top of the DNA-binding cleft, trapping the DNA and nucleotide in the active site. (C) Stereo diagrams comparing the open (magenta) and closed (yellow) forms of Klentaq1 bound to DNA and dNTP. (D) and (E) Stereo diagrams of the active site, showing the contacts between the polymerase (O helix), DNA, and ddCTP in the closed form and open form, respectively. Note the displacement of Tyr 671 from its stacking arrangement with template base G_{T5} in going from the open form (E) to the closed form (D), which allows the next template base (G_{T4}) to get in position for base pair formation with the ddCTP. Figure adapted from [62].

The structural data strongly suggest that the Klentaq1_DNA and **Klentaq1***_DNA_dNTP complexes respectively represent DNA polymerase conformations before binding dNTP and after the rate-limiting step, when the polymerase is predicted to clamp down on its substrates in preparation for chemistry (Fig. **2**). The presence of the intermediate open Klentaq1_DNA_dNTP complex implies that the enzyme can transition between binding dNTP, sensing it for "correctness" and releasing the wrong dNTP, before finally closing tightly around the right one. The new base pair must fit in a tight binding pocket (Fig. **4**; Fig. **5D**); therefore, the incompatible geometry of a mismatched base pair may trigger transition to an open conformation and dNTP release, while the correct fit of a Watson-Crick base pair allows the polymerase to proceed to the nucleotidyl transfer step.

The nucleotide binding site in DNA polymerases can also discriminate between dNTPs and rNTPs (ribonucleoside triposphates) such that the polymerase synthesizes DNA not RNA. For example, in the T7 DNA polymerase, the C-2 carbon of the ribose moiety fits in a tight pocket that cannot accommodate the 2'-hydroxyl group of a rNTP [37]. Other polymerases utilize a "gating" mechanism to sterically block the 2'-hydroxyl group of an incoming rNTP. In the Klenow fragment, for example, a glutamate residue in the nucleotide-binding site plays a key role in excluding the 2'-hydroxyl group

from the site [64]. A similar mechanism (using a phenylalanine residue) for blocking rNTP entry has been reported for the moloney murine leukemia virus reverse transcriptase [65]. Such intricate design features allow DNA polymerases to selectively synthesize DNA in spite of the 10-fold higher concentrations of rNTPs over dNTPs in the cell.

C) Processivity

The polymerase mechanism also predicts that after nucleotidyl transfer the polymerase switches back to an open conformation, from which the pyrophosphate and DNA products are released (Fig. **2**). In most polymerases, however, complete dissociation of DNA occurs very slowly (0.2 per second for T7 DNA polymerase [51]; Fig. **2**). Instead, these enzymes function processively by translocating on DNA and repeating the polymerization reaction, thus extending the primer by more than one nucleotide per template binding event [7]. Replicative DNA polymerases in particular are highly processive, and employ accessory protein factors to replicate thousands of nucleotides per DNA binding event. The T7 polymerase bound to thioredoxin exhibits a processivity of ~ 1500 base pairs at 20 _C (processivity = rate of polymerization / rate of DNA dissociation) [51,66,67]. Other DNA polymerases, such as Pol I [47] and HIV-RT [49], do not utilize accessory proteins and yet are capable of incorporating several nucleotides per DNA binding event. Thus, the inherent design of the DNA polymerase plays an important role in processivity. The crystal structures of DNA-bound polymerases, such as the T7 DNA polymerase [37] and Klenow fragment [68] among others, reveal that the thumb domain forms a cap-like structure over the DNA binding cleft and covers a significant portion of the duplex. Residues in the tip of the thumb contact the primer-template upto several bases away from the active site, and appear to increase the grip of the polymerase on DNA [37]. Consistent with the structural data, removal of a 24 amino acid domain from the tip of the Klenow fragment thumb substantially reduces its affinity for DNA and also reduces its processivity [69]. Similarly, mutations in the thumb domain of HIV-1 reverse transcriptase increase the DNA dissociation rate and thus reduce processivity [70]. Simply the tight binding of polymerase to DNA solves only half the problem, since processivity also requires that the polymerase release DNA and translocate at the end of each catalytic cycle. The structures of Klentaq1 bound to DNA in Fig. **5A** and **5B** show the DNA almost completely enclosed in a cylinder-like groove formed by the fingers and palm and capped by the thumb [62]. In the closed conformation, the DNA is wedged tightly within the active site and does not move. In the open conformation the DNA is still enclosed in the groove, but may have the space to may move freely without complete dissociation from the polymerase. Thus, in the open conformation the polymerase can directly translocate to the next base on the template for processive DNA replication.

D) DNA Editing

On the rare occasions when a mismatch is incorporated into DNA, the primer strand partitions to

the 3' – 5' exonuclease site on the polymerase where the mismatch is rapidly excised [19,34]. In the crystal structures determined so far, the polymerase and exonuclease active sites are separated by a distance of about 30 – 40 Å (e.g., Klenow fragment [4]; RB69 polymerase [39]. The DNA must move from the polymerase site and unwind a few base pairs to present the primer strand to the exonuclease site for editing. The structure of the RB69 polymerase_DNA editing complex shows how the DNA is partially melted and bent such that 3 – 4 bases of the primer enter a deep cleft, ending with the 3'-primer terminus in the active site (Fig. **6A**) [71]. The exonuclease active site contains two metal ions that are ligated by acidic residues, and function analogously to the ions in the polymerase active site. One ion helps deprotonate a water molecule to generate the hydroxide ion for nucleophilic attack on the phosphorus in the phosphodiester bond. Both ions coordinate the oxygens of the phosphodiester bond, thereby stabilizing the geometry and the charge of the transition-state, and facilitating exit of the oxyanion nucleotide product.

A terminal mismatched base pair in the primer-template (or even mismatches in the –2 or –3 position) inhibits the next reaction cycle by blocking the rate-limiting conformational change in the polymerase (thereby stalling the polymerase in the "open" conformation), and by drastically slowing the rate of chemistry (300 per second to 0.01 per second for the T7 DNA polymerase) [52]. During this pause in the reaction, the DNA is released from the polymerase site and directed to the exonuclease site for correction. Mismatched base pairs distort the DNA duplex, and it has been proposed that their incorrect geometry may trigger translocation of the primer strand to the exonuclease active site for editing [34,72]. The T7 DNA polymerase structure reveals conserved amino acids in the fingers and palm domains that contact the final base pairs in the primer-template, and are in position to sense distortions in the minor groove caused by mismatches [37,61]. Mutations of analogous amino acids in the Klenow fragment reduce its affinity for the primer-template DNA and also slow the rate of DNA synthesis [73,74]. These results are consistent with the idea that mismatch recognition *via* a geometry-sensing mechanism changes the interaction between DNA polymerase and the primer-template, and may trigger conformational changes that stall the polymerase and allow the DNA to switch to the exonuclease site.

A comparison of the RB69 polymerase_DNA editing complex with a model structure of DNA bound to the polymerase site shows how both the polymerase and DNA undergo changes during the switch from DNA synthesis mode to DNA editing mode (Fig. **6B**) [71]. The double-stranded portion of the primer-template changes its trajectory by a 30_ angle as the primer strand is unwound and inserted into the exonuclease site. In the polymerase, the thumb domain shifts from a "closed" conformation (in the apo form) to a more "open" conformation, and a β hairpin structure that normally blocks access to the exonuclease is also shifted, allowing the primer clear access to the exonuclease active site. Furthermore, contacts between the β hairpin and the primer appear to stabilize the melted duplex and likely aid the entry of single-stranded DNA into the exonuclease and

facilitate DNA editing [75-78]. Similarly, in the Klenow fragment, numerous hydrogen-bonding and hydrophobic contacts stabilize single-stranded DNA in the exonuclease site [55,79].

Fig. (6). DNA editing. (A) Structure of RB69 DNA polymerase with DNA bound to the editing site. The polymerase and exonuclease active sites are indicated by P and E, respectively, and are 38 Å apart. (B) Comparison of the RB69 polymerase in editing mode (exo) *versus* the polymerase mode (pol). The β hairpin partitioning the two active sites is highlighted. Figure adapted from [71].

Since single-stranded DNA is the preferred substrate of the exonuclease, the stability of the DNA duplex likely influences the transfer of DNA between the polymerase and exonuclease active sites (T4 DNA polymerase [53,80]; *E. coli* Pol III [81]; Klenow fragment [82,83]). Mismatches disrupt the hydrogen bonding between base pairs and increase the melting capacity of DNA, which in turn may increase the potential for DNA editing [84,85]. Biochemical studies examining the influence of base pair hydrogen bonding on polymerase and exonuclease activity utilize nucleoside analogs that mimic the shape of native dNTPs but lack the ability to form hydrogen bonds [86,87]. When paired with natural bases in the terminal base pair of a primed DNA template, these analogs mimic Watson-Crick base pair geometry but form unstable base pairs because of the lack of hydrogen bonding. A recent study shows that the weak T : Z (4-methylbenzimidazole) and A : F (2,4-difluorotoluene) base pairs are edited with similar efficiency as mismatched base pairs such as T : C or A : A (in spite of their resemblance to the correct A : T base pair structure), while the correct, hydrogen bonded A : T base pair is edited at a significantly slower rate by the Klenow fragment [88]. These results further substantiate the hypothesis that base pair instability (as in a mismatch) and consequent DNA melting stimulates exonuclease activity and therefore has a strong influence on the editing mechanism.

Once single-stranded DNA binds to the exonuclease site it can be hydrolyzed very rapidly, so the question arises as to how the enzyme prevents extensive and unecessary degradation of DNA. The T7 polymerase mechanism indicates that DNA can easily slide back and forth between the two active sites, and that binding of the primer-template to the polymerase active site is energetically favored over binding to the exonuclease [34,72]. In other words, the DNA has a very small lifetime at the exonuclease site compared to the polymerase site. Consequently, DNA excision is not very processive, and after a mismatch is corrected, the DNA can exit rapidly from the exonuclease. The polymerase site rebinds the corrected DNA with high affinity and returns to the task of DNA synthesis. Studies of the T4 [89], herpes simplex virus [90], and ϕ29 [91] polymerases suggest that these enzymes also allow DNA to slide freely between the two active sites. Whether this property maintains low exonuclease processivity and controlled DNA editing in other polymerases, as in the T7 polymerase, remains to be seen. Ongoing studies of different polymerases continue to clarify structural and kinetic parameters governing polymerase and exonuclease activity, and are anticipated to provide greater insights into the mechanisms by which polymerases replicate DNA with high

fidelity.

HOLOENZYME STRUCTURE AND FUNCTION

During genomic DNA replication, several different proteins function at the DNA replication fork as part of a replisome (Fig. 7) [7,8]. A DNA helicase unwinds duplex DNA to create a single-stranded template for the polymerase [92]. Single-stranded DNA binding protein (SSB) stabilizes the template strands [93], while primase synthesizes short RNA primers on DNA that serve as start sites for DNA replication [1]. The polymerase catalyzes DNA synthesis and proofreads the new DNA strand with its 3' - 5' exonuclease activity, as described above. Genomic DNA replication additionally requires that the DNA polymerase function with high speed and fidelity, as well as exceptionally high processivity. Moreover, the polymerase must also coordinate leading and lagging strand DNA replication, which is a challenging task given that the leading strand is synthesized continuously in the direction of the replication fork and the lagging strand must be synthesized discontinuously in the opposite direction, due to the anti-parallel nature of DNA [1]. Most polymerases acquire the specialized properties required for genomic DNA replication by associating with one or more accessory proteins. These proteins work together with the polymerase in a holoenzyme particle that is capable of simultaneous replication of leading and lagging strand DNA [94], at close to 1000 nucleotides per second [94], and with an error rate approaching 1 in 10^7 base pairs [95]. The following sections focus on the composition, architecture, and the mechanisms of action of DNA polymerase holoenzymes, using *E. coli* DNA polymerase III as a model system to understand the workings of these multi-protein machines in *E. coli* and other organisms.

Anatomy of a Holoenzyme

DNA polymerase holoenzymes generally contain three distinct components: a) a core DNA polymerase, b) a sliding clamp processivity factor, and c) a clamp loader. These components appear conserved in all cellular organisms ranging from bacteria to humans (Table I). The DNA polymerase III holoenzyme (Pol III) is made up of ten proteins, each of which contributes a different activity toward DNA replication (Fig. 7) [24]. The core DNA polymerase ($\alpha\varepsilon\theta$) comprises: α, the DNA polymerase, ε, the 3' – 5' proofreading exonuclease, and θ, whose function is unclear. Like most other polymerases, the core DNA polymerase has very low processivity and dissociates from the DNA after incorporating only about 10 to 20 nucleotides [96]. This problem is alleviated by the β protein, a sliding clamp that binds the polymerase with high affinity and increases its stability on template DNA [96,97]. Pol III also contains a 5-protein clamp loader, γ complex, that assembles the β processivity factor onto DNA for use by the core polymerase [98]. The tenth protein, τ, holds together two core polymerases, thus facilitating simultaneous synthesis of leading and lagging strands by the holoenzyme [99]. τ also binds the γ complex, serving as a connector that holds the polymerase and clamp loader in one holoenzyme particle [99]. Furthermore, τ interacts with the

DNA helicase (DnaB [100,101]) and the primase (DnaG [100]), stimulating their activity and helping coordinate holoenzyme function with other replisomal proteins at the DNA replication fork (Fig. 7).

Fig. (7). The DNA replisome. At the DNA replication fork, the helicase unwinds duplex DNA to make single-stranded template, which is stabilized by SSB, and primase synthesizes an RNA primer on the template for initiation of replication. The holoenzyme, comprising two core DNA polymerases (αεθ), a connector protein (τ), a clamp loader (γ complex), and two sliding clamps (β), replicates leading and lagging DNA strands simultaneously. (A) The primase maintains its grip on the new primer through an essential contact with SSB. (B) The χ subunit of γ complex severs the primase-SSB contact as the clamp loader takes its place on the primer and assembles the circular β clamp around it. The primase recycles to another site on the template to continue synthesizing primers. (C) When the lagging strand polymerase finishes an Okazaki fragment, it releases the clamp and DNA, and reassembles at the upstream primed site (with β on it) to start a new fragment. Figure adapted from [139].

Eukaryotic DNA polymerase holoenzymes have similar components, including Pol δ (and Pol ε?) the replicative DNA polymerase, PCNA, the sliding clamp, and RFC, the clamp loader [102]. So far, there appears to be no equivalent to the *E. coli* τ protein connector, although, the RFC subunits share significant sequence homology with τ [103], and at least one RFC subunit appears to bind Pol δ [104,105]. The three eukaryotic holoenzyme components may also interact with each other directly to form a stable holoenzyme particle. The *S. cerevisiae* Pol δ (a three subunit polymerase) forms a dimer, which is likely responsible for simultaneous replication of leading and lagging DNA strands [106]. Similarly, the four-subunit *Schizosaccharomyces pombe* DNA Pol δ forms a dimer, which may serve the same function [107]. The PCNA clamp is bound tightly by Pol δ during DNA replication, and studies with the human DNA polymerase indicate that RFC also associates with the holoenzyme through direct interaction with Pol δ [104,105], possibly helped by a contact with RPA (the eukaryotic SSB) on DNA [105]. The Cdc45 protein, which interacts with the eukaryotic replicative helicase MCM [108], and moves with the replication fork [108], may also be involved in organizing the eukaryotic replisome.

Table I. Components of the DNA Replisome

Processivity Factors

A) Circular Sliding Clamps

Defined as the ratio of the DNA polymerization rate *versus* the DNA dissociation rate, processivity reflects the ability of a polymerase to extend DNA per template binding event. Replicative DNA polymerases tend to be highly processive and can extend a new strand by several thousand nucleotides without falling off the template [1]. Studies of DNA replicases from various organisms have revealed different mechanisms by which polymerases acquire processivity. By far the most

common mechanism involves the use of "sliding clamps" as accessory factors. Sliding clamps are ring-shaped proteins that encircle double-stranded DNA and form a topological link with it [110]. This stable yet nonspecific, sequence-independent interaction with DNA allows the ring to slide freely along the duplex. These clamps also bind DNA polymerase, and thus serve as mobile tethers to keep the polymerase associated with the template during DNA synthesis.

The term "sliding clamp" was first used to describe the ability of bacteriophage T4 polymerase accessory factors (gp45 clamp in concert with gp44/62) to stimulate DNA replication [111-112]. Early biochemical studies with the *E. coli* Pol III holoenzyme suggested that one of its accessory subunits, β, is in fact the sliding clamp—a ring-shaped protein that encircles DNA [97]. The experiments showed that β assembled on a circular DNA molecule remains stably bound, but if the DNA is linearized with an endonuclease, the clamp dissociates rapidly and completely from the DNA [97]. It was hypothesized that the circular clamp slides freely on DNA and can therefore fall off the ends off a linear DNA molecule. This hypothesis was confirmed by the crystal structure of β, which revealed a protein ring about 80 Å in diameter with a central hole about 35 Å across (Fig. **8A**) [113]. Two β monomers interact in a head-to-tail fashion to form a tightly closed ring with no apparent breaks at the interfaces. The central hole is more than large enough to accommodate A- or B-form DNA, and positive electrostatic potential on the inside of the clamp complements the negatively charged phosphate backbone of DNA. The absence of specific DNA binding elements inside the ring suggests that β interacts with the double helix through non-specific, water-mediated interactions, and explains how it can slide freely on DNA [110,113].

Soon after the *E. coli* clamp structure was solved, the crystal structures of sliding clamps from *S. cerevisiae* (yPCNA) [114], humans (hPCNA) [115], bacteriophage T4 (gp45) [116], and bacteriophage RB69 [71] were determined (Fig. **8A**). The proteins form remarkably similar tertiary and quaternary structures despite the lack of significant similarity between their primary amino acid sequences (except for the T4 and RB69 clamps, which are closely related). All the clamps are highly symmetrical toroids, with similar dimensions and a central hole of about 30 - 35 Å in diameter. The most striking difference between the clamps is that PCNA and the T4 and RB69 monomer units are about two-thirds the size of β, therefore three monomers interact with each other to form the clamp instead of two as in β. All other structure-function aspects of the clamps appear closely related; the monomers form an unbroken ring that can encircle duplex DNA, and each monomer consists of two (PCNA, T4, RB69) or three (β) domains with repeating motifs that maintain the ring structure, as well as inter-domain connector loops that are important for interactions with other proteins. As in β, the inside channel of the rings is wide, positively charged, and lined with α helices that do not make any specific contacts with DNA, allowing the clamps to slide freely on the duplex behind the polymerase (Fig. **8A**).

Fig. (8). Sliding clamps and their dynamic interactions with DNA polymerase. (A) Front views of circular sliding

clamps from bacteriophage T4 (gp45), bacteriophage RB69, *E. coli* (β), *S. cerevisiae* (yPCNA), and humans (hPCNA). (B) A model structure of the RB69 DNA polymerase tethered to primer-template DNA by the sliding clamp. (C) and (D) Models illustrating how the *E. coli* DNA polymerase associated with its β sliding clamp might bypass a secondary structure on DNA by sliding over it (C) or remove it by a strand displacement mechanism (D). Figure adapted from [71,120].

Recently, the bacteriophage RB69 clamp structure was solved in complex with a carboxyl-terminal peptide from the RB69 DNA polymerase (Fig. **8B**). This 11-residue, C-terminal peptide is essential for interaction between the polymerase and clamp (in both phage RB69 and T4), and is likely the major point of contact between the two [117-119]. The structure of the DNA polymerase shows the C-terminus extending out from behind the main body of the enzyme, much like a tether that hooks the polymerase onto the sliding clamp (Fig. **8B**) [39,71]. Hydrophobic residues among the last seven amino acids of the tether form tight contacts with a small hydrophobic pocket on the sliding clamp, while the remaining residues extend away from the clamp toward the main body and are structurally more flexible. This type of connection may be of advantage when the polymerase encounters pause sites during DNA replication and has to adjust its position on DNA [120]. For example, following misincorporation of a nucleotide, the polymerase pauses while the DNA switches between the polymerase active site and editing site. A flexible tether between the polymerase and clamp could facilitate such movements while minimizing the danger of the polymerase dissociating from the DNA altogether. In a recent study, the β clamp, as well as human PCNA, were found capable of sliding over secondary structures such as single-stranded loops, small stem-loops, and bubbles in DNA [121]. Furthermore, when the Pol III holoenzyme encountered such small obstacles, it did not fall off DNA but rather traversed the obstacle (possibly by piggybacking on the clamp) and resumed synthesis at a downstream 3'-terminus (Fig. **8C**). Given secondary structures that were too large for the clamp to slide over, the polymerase could clear them by a strand-displacement mechanism or eventually dissociate from DNA (Fig. **8D**) [121]. *In vivo*, such mechanisms may avert complete dissassembly of the holoenzyme at every pause site and facilitate rapid DNA replication, while repair processes fix the anomalous structures in DNA at a later time. Although structural details of the interaction between β and the core DNA polymerase are not yet clear, a flexible connection, similar to that detected in phage RB69, might allow the polymerase to shift position on DNA as it attempts to traverse or displace obstacles in its path.

B) Monomeric Tethers

Several replicative DNA polymerases acquire processivity through means other than the circular sliding clamps described above. One such enzyme is the bacteriophage T7 DNA polymerase that co-opts a bacterial host protein, thioredoxin, to increase its processivity from a few nucleotides to more than a thousand nucleotides per template binding event [122,123]. Thioredoxin functions as a monomer and forms a 1 : 1 complex with the polymerase by interacting with its extended thumb

domain, as shown in Fig. **4** [37]. The thumb domain covers a substantial portion of the DNA binding cleft and stabilizes the polymerase on DNA. It has been hypothesized that thioredoxin bound to the thumb further extends the "cap" structure and converts the enzyme itself into a clamp, thus increasing its grip on DNA [61]. Although the thioredoxin molecule in the crystal structure does not cap the DNA binding cleft (Fig. **4**), this extended conformation may represent an open form of the enzyme, which is required for initial interaction with DNA and for DNA release after the reaction is complete. It is also possible that thioredoxin does not trap DNA within the polymerase, but rather binds DNA through non-specific electrostatic contacts, which allow it to slide on DNA and thus function as a mobile tether for the polymerase.

Prominent among other polymerases that utilize monomeric processivity factors are the eukaryotic mitochondrial DNA polymerase γ, which is composed of a large catalytic subunit and a smaller accessory subunit that increases its processivity [124,125], and the herpes simplex virus DNA polymerase, which utilizes the UL42 protein for high processivity [126]. A model structure of the DNA polymerase γ accessory subunit predicts that it shares some structural similarity with thioredoxin, but there is no direct evidence yet that polymerase γ employs the same mechanism for processivity as the T7 DNA polymerase [127]. The UL42 protein is even more of a mystery, since a recent report shows that its structure is very similar to that of a PCNA monomer (Fig. **8A**), despite the lack of any sequence homology between the two [128]. Furthermore, a C-terminal peptide of the HSV DNA polymerase is critical for binding to UL42 [129], suggesting that interaction between the polymerase and UL42 is through a flexible tether, much like that observed for the RB69 polymerase and clamp [71,128]. Yet unlike the RB69 clamp and PCNA, UL42 functions as a single monomer in a 1 : 1 complex with the HSV DNA polymerase [126]. Although it is not yet clear how the UL42 structure corresponds with its function as a processivity factor, it binds duplex DNA with high affinity [130,131], and may utilize electrostatic interactions to slide on DNA, as has been proposed for thioredoxin. The UL42 protein may reflect a precursor form of the circular sliding clamps as its structure strongly resembles that of a sliding clamp subunit but its action appears more primitive (without formation of a multimeric ring or topological linkage with DNA).

Differences in the structures of various processivity factors are mirrored by differences in their mechanisms for conferring processivity on DNA polymerases. As described earlier, processivity is a function of two factors: the propensity of the polymerase to extend DNA *versus* its propensity to dissociate from DNA. In keeping with the proposed capped structure of T7 DNA polymerase and thioredoxin on DNA, a study of T7 polymerase-catalyzed DNA synthesis indicates that thioredoxin causes a significant decrease in the dissociation of the polymerase from DNA [51,122]. However, thioredoxin also increases polymerase processivity by increasing the rate of DNA polymerization. It should be interesting to determine how an accessory protein bound so far from the catalytic site influences the rate of polymerization (Fig. **4**). UL42 also reduces the rate of HSV polymerase

dissociation from DNA, but appears to have no effect on the rate of polymerization [131]. In contrast, the 8-fold increase in overall processivity of human mitochondrial DNA polymerase γ in the presence of its accessory subunit is almost entirely due to a substantial increase in the polymerization rate [132]. The accessory subunit has no effect on the dissociation of γ polymerase from DNA, implying that it does not function by physically trapping the polymerase on DNA, as suggested for thioredoxin and the circular sliding clamps. The diversity in structures and mechanisms for processivity even among the few DNA replicases examined in detail suggests that several different methods have evolved to effectively meet the essential requirement of rapid genomic DNA replication by DNA replicases.

Circular Clamp Assembly on DNA

A) Clamp Loader Structure

The crystal structure of a sliding clamp provided a clear explanation of how it serves as mobile tether for the DNA polymerase. It also raised an intriguing question; how does the closed protein ring form a topological link with DNA? In order to serve as processivity factors for DNA polymerase, sliding clamps must encircle primed DNA sites that serve as starting points for DNA replication. These proteins cannot self-assemble around DNA and must be loaded onto DNA by a protein complex called the clamp loader (Fig. 7). The clamp loader functions as a molecular matchmaker, using energy from ATP binding and hydrolysis to bring together a clamp and primed DNA and form a topological link between the two [133]. DNA replicases from a variety of organisms ranging from bacteriophage to humans utilize such ATP-fueled molecular matchmakers for processivity (Table I).

In *E. coli*, early studies of the Pol III holoenzyme had shown that 2 ATP molecules are hydrolyzed during formation of an initiation complex in which the polymerase is tethered to primed DNA by the β sliding clamp [134]. More recently, the *E. coli* clamp loader, γ complex, was found to hydrolyze 2 - 3 ATP molecules for assembly of one β clamp on DNA, confirming that its ATPase activity is directly coupled to formation of a processive DNA polymerase [135]. The *E. coli* clamp loader is a complex machine composed of two to four γ, one δ, one δ', as well as one each of the χ and ψ subunits [98,136]. The three subunits, γ, δ, and δ', are essential for clamp assembly (Fig. **9A**) [137]; the χ subunit aids holoenzyme assembly onto primed DNA [138,139] (discussed in a following section), and the function of ψ is unknown, although it appears to stabilize the γ complex under high salt conditions [140]. The basis for clamp assembly lies in δ, which is the only subunit that binds β and is capable of opening the ring without help from the other subunits or a nucleotide cofactor [135]. However, when δ is assembled within γ complex, it has very low affinity for β and cannot open the clamp in the absence of ATP. Under these conditions, the β binding site on the δ subunit is buried within γ complex, likely occluded by the δ' subunit. In the presence of ATP, the γ complex undergoes a conformational change that allows δ to bind β [135,141]. The crystal structure of the δ'

subunit provides some clues as to the nature of this conformational change. δ' is a "C"-shaped protein with globular top and bottom domains connected by a flexible hinge domain [142]. The γ subunit shares extensive sequence homology with the δ' subunit, and a model structure of γ based on δ' indicates that γ is also "C"-shaped (Fig. **9B**). The γ subunit contains the Walker A and B ATP binding motifs, and is the only subunit in γ complex that binds and hydrolyzes ATP [143,144]. The ATP-binding site abuts the hinge domain in γ, predicting that ATP binding and hydrolysis affect the flexible hinge and drive movement of the top and bottom domains relative to each other [142]. Thus, γ is the "motor" subunit in γ complex, and clamp assembly begins when at least two γ subunits bind ATP, resulting in a change in γ complex conformation (Fig. **9A**) [145]. This conformational change moves the δ' subunit and exposes δ for interaction with β. A study of β clamps covalently cross-linked at one or both interfaces of the dimer indicates that opening of one interface is necessary and sufficient for loading β onto DNA. Thus, on ATP binding, the δ subunit likely opens the β clamp at one interface in preparation for clamp assembly (Fig. **9A**).

Prior to ATP hydrolysis, the ATP-bound clamp loader also binds primed DNA to form an intermediate complex in which the open clamp and DNA substrate are held in close proximity, possibly in position to form a topological link with each other [145]. Completion of clamp assembly requires ATP hydrolysis, as ATPγS (an ATP analog not hydrolyzed by γ complex) cannot substitute for ATP in clamp assembly. β and DNA binding experiments performed with ATP and ATPγS show that the ATP-bound form of γ complex has high affinity for β and DNA, but both substrates are released when ATP is hydrolyzed [145]. A minimal model pathway for clamp assembly suggests that once the clamp loader_ATP_open clamp_DNA intermediate is formed, ATP hydrolysis reverts γ complex conformation back to the original state where δ' once again binds δ and disrupts its interaction with β. Consequently, β and DNA are released from γ complex, topologically linked to each other [135,145]. Investigation of γ complex activity using rapid kinetic techniques indicates that coupling between the ATPase reaction and clamp assembly is a bit more complex than the simple two-step model mechanism outlined above [146]. The γ complex binds two ATP molecules rapidly, at an apparent bimolecular rate of 4×10^5 M^{-1} $second^{-1}$ (Fig. **9A**). Interaction with primed DNA also occurs rapidly, on the order of 10^8 M^{-1} $second^{-1}$ (measured with γ complex alone [147]) and may occur before or after interaction with β. Next, a slow, intramolecular step occurs at 12 – 15 per second, and this has been attributed to opening of the β ring. Following ring opening, the two ATP molecules are hydrolyzed sequentially by γ complex. Hydrolysis of one ATP occurs at > 80 $second^{-1}$ and is followed by a rate-limiting step in the reaction. Thus, the second ATP is hydrolyzed much slower than the first, at an apparent rate of 2 $second^{-1}$. Hydrolysis of one ATP molecule appears sufficient for release of β from the γ complex, although it is not clear yet whether β release occurs after the first or the second ATP is hydrolyzed [146]. The sequential mechanism is indicative of multi-step coupling of ATP hydrolysis to clamp assembly, and the initial evidence suggests that each hydrolysis step is coupled to a distinct functional event in the pathway. Once the clamp

loader_ATP_open clamp_DNA intermediate is formed, the clamp loader may need to position the clamp, or DNA, or both, such that the DNA lies in the center of the open clamp. Further rearrangements may be necessary to close the clamp around DNA, release the clamp and DNA, and for the γ complex to revert back to its original conformation. These events are coordinated and likely coupled in a step-wise fashion to sequential ATP hydrolysis and release of ADP and P_i from γ complex.

Clamp loaders from other organisms are also multi-component machines that utilize energy from ATP binding and hydrolysis to assemble circular clamps on DNA (Table I). The eukaryotic clamp loader, Replication Factor C (RFC), contains five subunits that are highly conserved from yeast to humans (*S. cerevisiae* RFC1 – 5; human p140, p37, p36, p40, p38) [8,148]. All five proteins also share high homology with the γ and δ' subunits of *E. coli*, the gp44 subunits of the bacteriophage T4 clamp loader, gp44/62, as well as RFC1 and RFC3, the two subunits of the archaeal clamp loader [23,103]. The sequence similarities imply that these clamp loader proteins have a 'C'-shaped structure like γ and δ', and might employ this flexible, ATP-sensitive shape like γ complex for the mechanics of opening the circular clamp and closing it around DNA. Another striking similarity between γ complex and clamp loaders from other organisms is that they all contain several ATPase-active subunits. The most highly conserved motifs in all the proteins listed above are the classic Walker A and B sequences, which form the phosphate-binding loop (P loop) and Mg^{2+} binding sites, respectively, and are required for ATP binding and hydrolysis (Fig. **9B**). At least four of the five human RFC proteins contribute to RFC ATPase activity, since mutations in their ATP-binding sites reduce PCNA loading activity and processive DNA replication [149]. Multiple yeast RFC subunits also exhibit ATP binding/ATPase activity [150,151]. The bacteriophage T4 clamp loader contains four copies of the gp44 subunit [152], and the archaeal clamp loader mthRFC, isolated from *Methanobacterium thermoautotrophicum* ΔH, contains several copies of the RF1 and RFC3 subunits [23]. This organization of multiple ATPase-active subunits in clamp loaders suggests that the sequential ATPase mechanism employed by γ complex may generalize to clamp loaders from other organisms.

The bacteriophage T4 clamp loader, gp44/62, catalyzes assembly of the circular gp45 clamp onto DNA in a reaction driven by binding and hydrolysis of four ATP molecules [153,154]. Like γ complex, gp44/62 binds the gp45 clamp on binding ATP [155,156]. The clamp must open wide for assembly around primed DNA, and different studies suggest that clamp opening occurs prior to ATP hydrolysis (as seen for γ complex and β) or concurrent with the hydrolysis of two ATP molecules [157,158]. The gp44/62_gp45 complex also binds DNA, after which hydrolysis of the other two ATP molecules appears to drive conformational changes in gp44/62 and close the gp45 ring around DNA [159]. This proposed mechanism is similar to that of γ complex in that the ATPase activity appears sequential (two ATP molecules instead of one are hydrolyzed in each step), but the manner in which

ATP hydrolysis is coupled to clamp assembly may be different. Further details on both gp44/62 and γ complex activities should facilitate better comparison of the two clamp loaders.

Fig. (9). Clamp loader structure and mechanism. (A) A model structure of γ, based on the δ' crystal structure, shows the "C"-shape of the protein with a top and bottom domain connected by a hinge domain. The ATP site in γ lies in close proximity to the hinge domain, implying that ATP binding and hydrolysis might induce global changes in the shape of the protein. (B) A model mechanism of ATP hydrolysis-coupled clamp assembly shows that at first two γ subunits in γ complex rapidly bind ATP. The ensuing change in γ complex conformation exposes δ, resulting in interaction with β and slow opening of the β ring at one dimer interface. The ATP-bound clamp loader also binds primer-template DNA rapidly. Next, sequential ATP hydrolysis followed by ADP and P_i dissociation drive the step-wise rearrangement and release of β and DNA, resulting in β linked topologically around DNA. The arrows indicate that changes in β or DNA may be coupled to the first or second ATP hydrolysis step. Figure adapted from [142,146].

In eukaryotes, ATP binding to RFC facilitates stable interaction between RFC, PCNA, and DNA, and ATP hydrolysis results in a PCNA_DNA complex that is used effectively by the polymerase for processive DNA replication [160,161]. Like the δ subunit in γ complex, the p40_p37_p36 subcomplex of human RFC, and possibly the p40 subunit alone, is capable of opening the PCNA ring in the absence of ATP [162]. Applying the γ complex model mechanism for clamp loading to RFC suggests that ATP binding might induce a change in RFC conformation that exposes the p40 or p40_p37_p36 subcomplex and allows it to bind and open PCNA. Subsequent ATP hydrolysis would then drive formation of a topological link between PCNA and DNA, release of PCNA_DNA, and turnover of RFC, although these steps have not yet been examined individually. In spite of the overall similarities between RFC and γ complex, their mechanisms may differ in detail. For example, in addition to the p40_p37_p36 subcomplex, the large p140 subunit also binds PCNA [163,164], and this interaction may be required for both ring opening (measured as removal of clamps from circular DNA) and clamp assembly [162]. Moreover, unlike *E. coli* δ, the p40_p37_p36 subcomplex also has a DNA-stimulated ATPase activity, which could be important for coupling ATP hydrolysis to assembly of the clamp around DNA [162,165]. The role of each subunit (and its ATPase activity) in the mechanics of PCNA loading is still under investigation. Extensive genetic and *in vitro* studies of the yeast and human clamp loader proteins also suggest that RFC is involved in cell cycle cycle checkpoint pathways [148]. For example, the large RFC subunit (p140) can be phosphorylated by protein kinases such as PK II, and the phosphorylated form is less active in DNA replication [166]. This has led to speculation that the clamp loading activity of RFC may be modulated by mechanisms that control progression of the cell cycle. Site-specific cleavage of p140 by proteases such as caspase-3 has also been detected, implicating RFC activity in apoptosis-linked pathways [167,168]. Future studies are anticipated to not only clarify the mechanism of RFC-catalyzed clamp assembly but also shed light on how RFC activity is linked to cell cycle control mechanisms.

Holoenzyme Assembly on DNA

Clamp assembly on DNA is only one event in an ordered series of reactions that occur during assembly of the holoenzyme on DNA and the start of DNA replication (Fig. 7). In the beginning, the DNA duplex is unwound by a helicase [92] and single-stranded (ss) DNA binding protein (SSB) stabilizes the resulting ss-DNA templates [93]. Recent studies have revealed that SSB plays a far more complex role in DNA replication than simply binding to ss-DNA [138,139]. In fact, in both bacterial (*E. coli*) and eukaryotic (human) DNA replication systems, it plays a central role in coordinating the interactions between various proteins during assembly of a functional polymerase holoenzyme on DNA [105,139].

The primase (*E. coli* DnaG) targets to single-stranded DNA, possibly aided by interaction with the DnaB helicase [169] and synthesizes small RNA primers up to 12 nucleotides in length (Fig. **7A**). The primase remains tightly associated with the primer on DNA, stabilizing the small primer on DNA and protecting it against nucleases [139]. Next, the primase is displaced from the primer by the clamp loader (γ complex), which has to assemble the β clamp on primed DNA (Fig. **7B**). This switching of proteins at the primed DNA site occurs through a competition between the primase and the clamp loader for the SSB protein. After synthesizing the RNA primer, the primase remains on the RNA_DNA aided by stable interaction with SSB coated on the adjacent ss-DNA (Fig. **7A**). A single subunit of the clamp loader (χ) also binds SSB and by virtue of this interaction it destabilizes the primase-SSB complex, resulting in displacement of the primase from DNA. The released primase can now recycle to another site on ss-DNA, while the γ complex proceeds to load β at the primed site (Fig. **7B**) [139]. Clamp assembly is followed by another protein switch between the γ complex and the polymerase [170]. Both proteins compete directly for the same binding site on the β clamp. In solution, γ complex binds the clamp with higher affinity than the polymerase, but releases the clamp after loading it on primed DNA. In contrast, the polymerase has weak affinity for β in solution, but binds the clamp tightly once it is loaded on DNA. Thus, after γ complex completes β assembly on DNA and leaves, the polymerase can finally access the primer and clamp and initiate DNA replication [170].

Similar "protein switching" mechanisms for polymerase assembly have been discovered in the eukaryotic DNA replisome as well [102,105]. In the human system, polymerase/primase α synthesizes an RNA primer of about 35 - 50 bases on a single-stranded DNA template [171]. As observed in *E. coli*, Pol α remains firmly attached to the primed DNA site through interaction with RPA, the eukaryotic single-stranded DNA binding protein. In the first protein switch, the clamp loader, RFC, binds RPA and disrupts the RPA-Pol α interaction, resulting in displacement of Pol α from the DNA. Subsequently, RFC assembles a PCNA clamp on the primed DNA, but unlike γ complex, it remains bound to RPA and the PCNA_DNA complex [105]. In the second switch, the

replicative DNA polymerase δ competes with RFC for PCNA (Pol δ and RFC bind the same site on PCNA [172]) and RPA, and gains access to PCNA_DNA to complete holoenzyme assembly and initiate DNA synthesis. As in *E. coli*, the continuous presence of proteins on the primer (first Pol α, then RFC, and then Pol δ) protects it from degradation by nucleases. The pol α to RFC to Pol δ switch may also impose specificity to the replication reaction, as polymerases other than Pol δ (e.g., Pol I tested in this study) may not be capable of extending a primer bound by RFC [105].

Holoenzyme assembly occurs infrequently on the leading DNA strand since it is replicated in a continuous fashion (Fig. 7). On the lagging strand, however, the polymerase must finish an Okazaki fragment and restart a new fragment once every second or so [1,7]. When the lagging strand polymerase hits the end of an Okazaki fragment, it releases both the sliding clamp and DNA and has to be reassembled at an upstream primed DNA site as observed, for example, with the DNA polymerase III holoenzyme [100] (Fig. **7B**). As described above, holoenzyme assembly involves a series of reactions with changing protein-protein and protein-DNA interactions until the polymerase is finally in place on DNA. Completion of one reaction sets up another reaction, apparently because the next protein component has high affinity for the product of the previous reaction. The ordered nature of these reactions helps coordinate rapid cycling of the lagging DNA polymerase and thus ensures smooth and timely progression of the DNA replication fork.

General Application of the Principles of Holoenzyme Assembly

Protein switching mechanisms may operate in other situations during DNA replication, and possibly even in other DNA metabolic processes where several different proteins compete for one site of action. One likely candidate for a central role in protein switching mechanisms is the circular sliding clamp. Several enzymes, including the replication- and repair- specific endonucleases, Fen1 [173] and XPG [174], respectively, DNA ligase I (that seals nicks in DNA) [175], and DNA cytosine-5 methyl transferase (that methylates newly synthesized DNA) [176], bind the eukaryotic clamp, PCNA. All these enzymes contain the PCNA binding motif (a small, conserved sequence of hydrophobic amino acids similar to the RB69 polymerase C-terminal peptide that binds the clamp; Fig. **8B**), and presumably compete with each other and with DNA polymerase for binding to PCNA [177,178]. Ordered competition for the sliding clamp may impart a temporal sequence to the assembly and action of these enzymes in various DNA processing reactions. For example, Fen1 (flap endonuclease) processes Okazaki fragments by cleaving the 5' flap formed when DNA polymerase invades a downstream fragment by strand displacement synthesis [179]. If Fen1 competes successfully with the polymerase for PCNA in this situation, it could swiftly snip 5' flap structures before strand displacement proceeds too far. The PCNA binding motif is also found on the C-terminus of $p21^{CIP1}$, a cell cycle regulator that binds human PCNA and inhibits DNA replication [102,115,180]. Thus, PCNA may also play a central role in ordered reactions that link DNA replication with cell cycle control mechanisms.

In a DNA repair process such as nucleotide excision repair (NER) and mismatch repair, several proteins localize to a lesion (or mismatch) in DNA, excise the DNA containing the lesion, and then fill in the gap by DNA synthesis. In *E. coli*, the process of NER involves interactions between UvrA, UvrB, UvrC proteins, and DNA for ordered assembly and activity of the Uvr(A)BC excinuclease on DNA [11] UvrA functions as a molecular matchmaker to load UvrB on DNA, and then dissociates to allows UvrC into the complex. After excision, UvrC must leave before polymerase I can access the 3' - terminus and re-synthesize DNA across the gap. Furthermore, nucleotide excision repair is also coupled to transcription through the transcription-repair coupling factor (TCRF). TCRF can recognize RNA polymerase stalled at a lesion in DNA, facilitate disassembly of the transcription complex, and recruit the Uvr(A)BC excinuclease in a highly concerted reaction. Even more proteins function together in eukaryotic excision repair systems, including the DNA replication proteins, RPA, RFC, and PCNA [11]. In DNA mismatch repair, as in NER, several different proteins are involved in recognizing and fixing mismatched basepairs and insertion/deletion loops in DNA [16,17]. The highly coordinated activity required of proteins in each DNA repair pathway, and between the processes of DNA repair and replication (and recombination), implies that competitive protein switching mechanisms analogous to those observed in holoenzyme assembly may be employed in several DNA metabolic reactions.

NEW FRONTIERS

New technologies for crystal structure determination and rapid kinetic analysis have revealed intricate details of the design and mechanism of action of DNA polymerases. *In vivo* these enzymes often function in the context of multi-component assemblies in which accessory proteins influence polymerase activity as, for example, in DNA polymerase holoenzymes at the replication fork. Purification of these proteins from different and complex organisms, including yeast and humans, as well as reconstitution of holoenzymes *in vitro* is proceeding rapidly at this time. These advances, accompanied by an understanding of holoenzyme biochemistry, are laying the foundation for detailed structural and kinetic studies which are anticipated to further clarify how these large and complex biological machines work on DNA.

It should be noted that interest in archaeal DNA replication has intensified in the past few years, in part due to the complete sequencing of several archaeal genomes. The DNA replication machinery in these organisms is by and large analogous to that in eukaryotes, except that archaeal proteins appear less complex in structure and composition [23]. Consequently, these proteins present attractive model systems for studies of holoenzyme / replisome structure and function, such as outlined above. An added advantage provided by these systems is the opportunity to understand how enzymes responsible for DNA replication and repair have adapted to extreme conditions.

The newly identified lesion bypass polymerases present yet another frontier in DNA polymerase

research. These enzymes function in DNA replication for error-free or error-prone bypass of DNA segments damaged by ultraviolet radiation or chemicals [910,28]. In *S. cerevisiae*, for example, when the DNA replicase stalls at a *cis-syn* thymine-thymine dimer (which distorts the DNA template), Pol η can access the primer-terminus and insert two adenines opposite the T-T dimer [181-183]. The enzyme has low processivity and dissociates rapidly from DNA, allowing the replicase to rebind DNA after lesion bypass and restart DNA replication. Although Pol η appears to recognize an A : T base pair, it functions with low fidelity compared to replicative DNA polymerases, implying it may be less sensitive to distorted DNA in the active site. Other lesion bypass polymerases are more promiscuous (e.g., *S. cerevisiae* Rev1/Pol ζ), and will insert any base opposite the T-T dimer, leading to error-prone lesion bypass [184]. Thus, even among this new and relatively small group of enzymes there are intriguing signs of variations on the polymerase theme. It will be interesting to determine how lesion bypass polymerases are designed for polymerization activity that contrasts dramatically with that of replicative DNA polymerases.

In summary, on the simple foundation of a two-metal ion mechanism for catalyzing phosphodiester bond formation, a vast array of DNA polymerases has evolved, facilitating not only the development of complex organisms, but also providing those complex organisms with the opportunity for continually exciting research in this field.

ACKNOWLEDGEMENTS

We are grateful to David Jeruzalmi for his invaluable assistance in preparing the figures, and to Dr. Gabriel Waksman and Dr. Yousif Shamoo for providing the images of RB69 and Klentaq1 polymerases, respectively. We also thank Dr. Megan Davey and Dr. Smita Patel for helpful discussions. This work was supported by NIH grant GM 38839 (M.O'Donnell).

REFERENCES

[1] Kornberg, A.; Baker, T. A. *DNA Replication 2nd Ed.*, W. H. Freeman: New York, **1992**.

[2] Kornberg, A.; Lehman, I. R.; Bessman, M. J.; Simms, E. S. *Biochim. Biophys. Acta,* **1989**, 57-8.

[3] Kornberg, A. *Science,* **1960**, *131*, 1503.

[4] Ollis, D. L.; Brick, P.; Hamlin, R.; Xuong, N. G.; Steitz, T. A. *Nature,* **1985**, *313*, 762-6.

[5] Steitz, T. A. *J. Biol. Chem.,* **1999**, *274*, 17395-8.

[6] Hubscher, U.; Nasheuer, H. P.; Syvaoja, J. E. *Trends Biochem. Sci.,* **2000**, *25*, 143-7.

[7] Baker, T. A.; Bell, S. P. *Cell,* **1998**, *92*, 295-305.

[8] Stillman, B. *Cell,* **1994**, *78*, 725-8.

[9] Goodman, M. F. *Trends Biochem. Sci.,* **2000**, *25*, 189-195.

[10] Johnson, R. E.; Washington, M. T.; Prakash, S.; Prakash, L. *Proc. Natl. Acad. Sci. USA,*

1999, *96*, 12224-6.

[11] Sancar, A. *Annu. Rev. Biochem.*, **1996**, *65*, 43-81.

[12] Mol, C. D.; Parikh, S. S.; Putnam, C. D.; Lo, T. P.; Tainer, J. A. *Annu. Rev. Biophys. Biomol. Struct.*, **1999**, *28*, 101-28.

[13] Parikh, S. S.; Mol, C. D.; Hosfield, D. J.; Tainer, J. A. *Curr. Opin. Struct. Biol.*, **1999**, *9*, 37-47.

[14] Scharer, O. D.; Deng, L.; Verdine, G. L. *Curr. Opin. Chem. Biol.*, **1997**, *1*, 526-31.

[15] Roberts, R. J.; Cheng, X. *Annu. Rev. Biochem.*, **1998**, *67*, 181-98.

[16] Modrich, P.; Lahue, R. *Annu. Rev. Biochem.*, **1996**, *65*, 101-33.

[17] Kolodner, R. D.; Marsischky, G. T. *Curr. Opin. Genet. Dev.*, **1999**, *9*, 89-96.

[18] Delarue, M.; Poch, O.; Tordo, N.; Moras, D.; Argos, P. *Protein. Eng.*, **1990**, *3*, 461-7.

[19] Joyce, C. M.; Steitz, T. A. *Annu. Rev. Biochem.*, **1994**, *63*, 777-822.

[20] Fuller, C. W.; Richardson, C. C. *J. Biol. Chem.*, **1985**, *260*, 3197-206.

[21] Karam, J. D.; Konigsberg, W. H. *Prog. Nucleic Acid Res. Mol. Biol.*, **2000**, *64*, 65-96.

[22] Cann, I. K.; Ishino, Y. *Genetics*, **1999**, *152*, 1249-67.

[23] Kelman, Z. *Current Protein and Peptide Science*, **2000**, 1, 1-25.

[24] Kelman, Z.; O'Donnell, M. *Annu. Rev. Biochem.*, **1995**, *64*, 171-200.

[25] O'Reilly, M.; Teichmann, S. A.; Rhodes, D. *Curr. Opin. Struct. Biol.*, **1999**, *9*, 56-65.

[26] Tang, M.; Shen, X.; Frank, E. G.; O'Donnell, M.; Woodgate, R.; Goodman, M. F. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 8919-24.

[27] Wagner, J.; Gruz, P.; Kim, S. R.; Yamada, M.; Matsui, K.; Fuchs, R. P.; Nohmi, T. *Mol. Cell*, **1999**, *4*, 281-6.

[28] Friedberg, E. C.; Gerlach, V. L. *Cell*, **1999**, *98*, 413-6.

[29] Goodman, M. F.; Tippin, B. *Curr. Opin. Genet. Dev.*, **2000**, *10*, 162-8.

[30] Uemori, T.; Sato, Y.; Kato, I.; Doi, H.; Ishino, Y. *Genes Cells,*,**1997**, *2*, 499-512.

[31] Loeb, L. A.; Kunkel, T. A. *Annu. Rev. Biochem.*, **1982**, *51*, 429-57.

[32] Echols, H.; Goodman, M. F. *Annu. Rev. Biochem.*, **1991**, *60*, 477-511.

[33] Kunkel, T. A. *J. Biol. Chem.*, **1992**, *267*, 18251-4.

[34] Johnson, K. A. *Annu. Rev. Biochem.*, **1993**, *62*, 685-713.

[35] Bloom, L. B.; Chen, X.; Fygenson, D. K.; Turner, J.; O'Donnell, M.; Goodman, M. F. *J. Biol. Chem.*, **1997**, *272*, 27919-30.

[36] Kim, Y.; Eom, S. H.; Wang, J.; Lee, D. S.; Suh, S. W.; Steitz, T. A. *Nature*, **1995**, *376*, 612-6.

[37] Doublie, S.; Tabor, S.; Long, A. M.; Richardson, C. C.; Ellenberger, T. *Nature*, **1998**, *391*, 251-8.

[38] Kiefer, J. R.; Mao, C.; Braman, J. C.; Beese, L. S. *Nature*, **1998**, *391*, 304-7.

[39] Wang, J.; Sattar, A. K.; Wang, C. C.; Karam, J. D.; Konigsberg, W. H.; Steitz, T. A.

Cell, **1997**, *89*, 1087-99.

[40] Zhao, Y.; Jeruzalmi, D.; Moarefi, I.; Leighton, L.; Lasken, R.; Kuriyan, J. *Structure Fold Des.,,* **1999**, *7*, 1189-99.

[41] Pelletier, H.; Sawaya, M. R.; Kumar, A.; Wilson, S. H.; Kraut, J. *Science,* **1994**, *264*, 1891-903.

[42] Kohlstaedt, L. A.; Wang, J.; Friedman, J. M.; Rice, P. A.; Steitz, T. A. *Science,* **1992**, *256*, 1783-90.

[43] Brautigam, C. A.; Steitz, T. A. *Curr. Opin. Struct. Biol.,* **1998**, *8*, 54-63.

[44] Jager, J.; Pata, J. D. *Curr. Opin. Struct. Biol.,* **1999**, *9*, 21-8.

[45] Bryant, F. R.; Johnson, K. A.; Benkovic, S. J. *Biochemistry,* **1983**, *22*, 3537-46.

[46] Dahlberg, M. E.; Benkovic, S. J. *Biochemistry,* **1991**, *30*, 4835-43.

[47] Kuchta, R. D.; Mizrahi, V.; Benkovic, P. A.; Johnson, K. A.; Benkovic, S. J. *Biochemistry,* **1987**, *26*, 8410-7.

[48] Kuchta, R. D.; Benkovic, P.; Benkovic, S. J. *Biochemistry,* **1988**, *27*, 6716-25.

[49] Kati, W. M.; Johnson, K. A.; Jerva, L. F.; Anderson, K. S. *J. Biol. Chem.,* **1992**, *267*, 25988-97.

[50] Frey, M. W.; Sowers, L. C.; Millar, D. P.; Benkovic, S. J. *Biochemistry,* **1995**, *34*, 9185-92.

[51] Patel, S. S.; Wong, I.; Johnson, K. A. *Biochemistry,* **1991**, *30*, 511-25.

[52] Wong, I.; Patel, S. S.; Johnson, K. A. *Biochemistry,* **1991**, *30*, 526-37.

[53] Capson, T. L.; Peliska, J. A.; Kaboord, B. F.; Frey, M. W.; Lively, C.; Dahlberg, M.; Benkovic, S. J. *Biochemistry,* **1992**, *31*, 10984-94.

[54] Dzantiev, L.; Romano, L. J. *Biochemistry,* **2000**, *39*, 356-61.

[55] Beese, L. S.; Steitz, T. A. *Embo J.,* **1991**, *10*, 25-33.

[56] Sawaya, M. R.; Pelletier, H.; Kumar, A.; Wilson, S. H.; Kraut, J. *Science,* **1994**, *264*, 1930-5.

[57] Huang, H.; Chopra, R.; Verdine, G. L.; Harrison, S. C. *Science,* **1998**, *282*, 1669-75.

[58] Moran, S.; Ren, R. X.; Kool, E. T. *Proc. Natl. Acad. Sci. USA,* **1997**, *94*, 10506-11.

[59] Morales, J. C.; Kool, E. T. *Nat Struct Biol,* **1998**, *5*, 950-4.

[60] Sawaya, M. R.; Prasad, R.; Wilson, S. H.; Kraut, J.; Pelletier, H. *Biochemistry,* **1997**, *36*, 11205-15.

[61] Doublie, S.; Ellenberger, T. *Curr. Opin. Struct. Biol.,* **1998**, *8*, 704-12.

[62] Li, Y.; Korolev, S.; Waksman, G. *Embo J.,* **1998**, *17*, 7514-25.

[63] Korolev, S.; Nayal, M.; Barnes, W. M.; Di Cera, E.; Waksman, G. *Proc. Natl. Acad. Sci. USA,* **1995**, *92*, 9264-8.

[64] Astatke, M.; Ng, K.; Grindley, N. D.; Joyce, C. M. *Proc. Natl. Acad. Sci. USA,* **1998**, *95*, 3402-7.

[65] Gao, G.; Orlova, M.; Georgiadis, M. M.; Hendrickson, W. A.; Goff, S. P. *Proc. Natl. Acad. Sci. USA,* **1997**, *94*, 407-11.

[66] Tabor, S.; Huber, H. E.; Richardson, C. C. *J. Biol. Chem.*, **1987**, *262*, 16212-23.

[67] Richardson, C. C. *Cell*, **1983**, *33*, 315-7.

[68] Beese, L. S.; Derbyshire, V.; Steitz, T. A. *Science*, **1993**, *260*, 352-5.

[69] Minnick, D. T.; Astatke, M.; Joyce, C. M.; Kunkel, T. A. *J. Biol. Chem.*, **1996**, *271*, 24954-61.

[70] Beard, W. A.; Minnick, D. T.; Wade, C. L.; Prasad, R.; Won, R. L.; Kumar, A.; Kunkel, T. A.; Wilson, S. H. *J. Biol. Chem.*, **1996**, *271*, 12213-20.

[71] Shamoo, Y.; Steitz, T. A. *Cell*, **1999**, *99*, 155-66.

[72] Donlin, M. J.; Patel, S. S.; Johnson, K. A. *Biochemistry*, **1991**, *30*, 538-46.

[73] Polesky, A. H.; Steitz, T. A.; Grindley, N. D.; Joyce, C. M. *J. Biol. Chem.*, **1990**, *265*, 14579-91.

[74] Polesky, A. H.; Dahlberg, M. E.; Benkovic, S. J.; Grindley, N. D.; Joyce, C. M. *J. Biol. Chem.*, **1992**, *267*, 8417-28.

[75] Spacciapoli, P.; Nossal, N. G. *J. Biol. Chem.*, **1994**, *269*, 438-46.

[76] Stocki, S. A.; Nonay, R. L.; Reha-Krantz, L. J. J. Mol. Biol., **1995**, *254*, 15-28.

[77] Baker, R. P.; Reha-Krantz, L. J. *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 3507-12.

[78] Wu, P.; Nossal, N.; Benkovic, S. J. *Biochemistry*, **1998**, *37*, 14748-55.

[79] Freemont, P. S.; Friedman, J. M.; Beese, L. S.; Sanderson, M. R.; Steitz, T. A. *Proc. Natl. Acad. Sci. USA*, **1988**, *85*, 8924-8.

[80] Clayton, L. K.; Goodman, M. F.; Branscomb, E. W.; Galas, D. J. *J. Biol. Chem.*, **1979**, *254*, 1902-12.

[81] Brenowitz, S.; Kwack, S.; Goodman, M. F.; O'Donnell, M.; Echols, H. *J. Biol. Chem.*, **1991**, *266*, 7888-92.

[82] Brutlag, D.; Kornberg, A. *J. Biol. Chem.*, **1972**, *247*, 241-8.

[83] Hochstrasser, R. A.; Carver, T. E.; Sowers, L. C.; Millar, D. P. *Biochemistry*, **1994**, *33*, 11971-9.

[84] Carver, T. E., Jr.; Hochstrasser, R. A.; Millar, D. P. *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 10670-4.

[85] Bloom, L. B.; Otto, M. R.; Eritja, R.; Reha-Krantz, L. J.; Goodman, M. F.; Beechem, J. M. *Biochemistry*, **1994**, *33*, 7576-86.

[86] Schweitzer, B. A.; Kool, E. T. *J. Org. Chem.*, **1994**, *59*, 7238-42.

[87] Guckian, K. M.; Morales, J. C.; Kool, E. T. *J. Org. Chem.*, **1998**, *63*, 9652-56.

[88] Morales, J. C.; Kool, E. T. *Biochemistry*, **2000**, *39*, 2626-32.

[89] Reddy, M. K.; Weitzel, S. E.; von Hippel, P. H. *J. Biol. Chem.*, **1992**, *267*, 14157-66.

[90] Strick, R.; Knopf, C. W. *Biochim. Biophys. Acta*, **1998**, *1388*, 315-24.

[91] de Vega, M.; Blanco, L.; Salas, M. *J. Mol. Biol.*, **1999**, *292*, 39-51.

[92] Patel, S. S.; Picha, K. M. *Annu. Rev. Biochem.*, **2000**.

[93] Lohman, T. M.; Ferrari, M. E. *Annu. Rev. Biochem.*, **1994**, *63*, 527-70.

[94] Studwell, P. S.; O'Donnell, M. *J. Biol. Chem.*, **1990**, *265*, 1171-8.

[95] Bloom, L. B.; Chen, X.; Fygenson, D. K.; Turner, J.; O'Donnell, M.; Goodman, M. F. *J. Biol. Chem.*, **1997**, *272*, 27919-30.

[96] Fay, P. J.; Johanson, K. O.; McHenry, C. S.; Bambara, R. A. *J. Biol. Chem.*, **1981**, *256*, 976-83.

[97] Stukenberg, P. T.; Studwell-Vaughan, P. S.; O'Donnell, M. *J. Biol. Chem.*, **1991**, *266*, 11328-34.

[98] Onrust, R.; Finkelstein, J.; Naktinis, V.; Turner, J.; Fang, L.; O'Donnell, M. *J. Biol. Chem.*, **1995**, *270*, 13348-57.

[99] Onrust, R.; Finkelstein, J.; Turner, J.; Naktinis, V.; O'Donnell, M. *J. Biol. Chem.*, **1995**, *270*, 13366-77.

[100] Yuzhakov, A.; Turner, J.; O'Donnell, M. *Cell*, **1996**, *86*, 877-86.

[101] Kim, S.; Dallmann, H. G.; McHenry, C. S.; Marians, K. J. *Cell*, **1996**, *84*, 643-50.

[102] Waga, S.; Stillman, B. *Annu. Rev. Biochem.*, **1998**, *67*, 721-51.

[103] O'Donnell, M.; Onrust, R.; Dean, F. B.; Chen, M.; Hurwitz, J. *Nucleic Acids Res.*, **1993**, *21*, 1-3.

[104] Pan, Z. Q.; Chen, M.; Hurwitz, J. *Proc. Natl. Acad. Sci. USA*, **1993**, *90*, 6-10.

[105] Yuzhakov, A.; Kelman, Z.; Hurwitz, J.; O'Donnell, M. *Embo J.*, **1999**, *18*, 6189-99.

[106] Burgers, P. M.; Gerik, K. J. *J. Biol. Chem.*, **1998**, *273*, 19756-62.

[107] Zuo, S.; Bermudez, V.; Zhang, G.; Kelman, Z.; Hurwitz, J. *J. Biol. Chem.*, **2000**, *275*, 5153-62.

[108] Dutta, A.; Bell, S. P. *Annu. Rev. Cell Dev. Bio.l*, **1997**, *13*, 293-332.

[109] Aparicio, O. M.; Weinstein, D. M.; Bell, S. P. *Cell*, **1997**, *91*, 59-69.

[110] Kuriyan, J.; O'Donnell, M. *J. Mol. Biol.*, **1993**, *234*, 915-25.

[111] Venkatesan, M.; Nossal, N. G. *J. Biol. Chem.*, **1982**, *257*, 12435-43.

[112] Huang, C. C.; Hearst, J. E.; Alberts, B. M. *J. Biol. Chem.*, **1981**, *256*, 4087-94.

[113] Kong, X. P.; Onrust, R.; O'Donnell, M.; Kuriyan, J. *Cell*, **1992**, *69*, 425-37.

[114] Krishna, T. S.; Kong, X. P.; Gary, S.; Burgers, P. M.; Kuriyan, J. *Cell*, **1994**, *79*, 1233-43.

[115] Gulbis, J. M.; Kelman, Z.; Hurwitz, J.; O'Donnell, M.; Kuriyan, J. *Cell*, **1996**, *87*, 297-306.

[116] Moarefi, I.; Jeruzalmi, D.; Turner, J.; O'Donnell, M.; Kuriyan, J. *J. Mol. Biol.*, **2000**, *296*, 1215-23.

[117] Alley, S. C.; Jones, A. D.; Soumillion, P.; Benkovic, S. J. *J. Biol. Chem.*, **1999**, *274*, 24485-9.

[118] Berdis, A. J.; Soumillion, P.; Benkovic, S. J. *Proc. Natl. Acad. Sci. USA*, **1996**, *93*, 12822-7.

[119] Goodrich, L. D.; Lin, T. C.; Spicer, E. K.; Jones, C.; Konigsberg, W. H. *Biochemistry*, **1997**, *36*, 10474-81.

[120] Hingorani, M. M.; O'Donnell, M. *Curr. Biol.*, **2000**, *10*, R25-9.

[121] Yao, N.; Hurwitz, J.; O'Donnell, M. *J. Biol. Chem.*, **2000**, *275*, 1421-32.

[122] Tabor, S.; Huber, H. E.; Richardson, C. C. *J. Biol. Chem.*, **1987**, *262*, 16212-23.

[123] Huber, H. E.; Tabor, S.; Richardson, C. C. *J. Biol. Chem.*, **1987**, *262*, 16224-32.

[124] Wang, Y.; Farr, C. L.; Kaguni, L. S. *J. Biol. Chem.*, **1997**, *272*, 13640-6.

[125] Carrodeguas, J. A.; Kobayashi, R.; Lim, S. E.; Copeland, W. C.; Bogenhagen, D. F. *Mol. Cell Biol.*, **1999**, *19*, 4039-46.

[126] Gottlieb, J.; Marcy, A. I.; Coen, D. M.; Challberg, M. D. *J. Virol.*, **1990**, *64*, 5976-87.

[127] Fan, L.; Sanschagrin, P. C.; Kaguni, L. S.; Kuhn, L. A. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 9527-32.

[128] Zuccola, H. J.; Filman, D. J.; Coen, D. M.; Hogle, J. M. *Molecular Cell*, **2000**, *5*, 267-78.

[129] Digard, P.; Bebrin, W. R.; Weisshart, K.; Coen, D. M. *J. Virol.*, **1993**, *67*, 398-406.

[130] Gallo, M. L.; Jackwood, D. H.; Murphy, M.; Marsden, H. S.; Parris, D. S. *J. Virol.*, **1988**, *62*, 2874-83.

[131] Weisshart, K.; Chow, C. S.; Coen, D. M. *J. Virol.*, **1999**, *73*, 55-66.

[132] Johnson, A. A.; Tsai, Y.; Graves, S. W.; Johnson, K. A. *Biochemistry*, **2000**, *39*, 1702-8.

[133] Sancar, A.; Hearst, J. E. *Science*, **1993**, *259*, 1415-20.

[134] Burgers, P. M.; Kornberg, A. *J. Biol. Chem.*, **1982**, *257*, 11468-73.

[135] Turner, J.; Hingorani, M. M.; Kelman, Z.; O'Donnell, M. *Embo J.*, **1999**, *18*, 771-83.

[136] Maki, S.; Kornberg, A. *J. Biol. Chem.*, **1988**, *263*, 6547-54.

[137] Onrust, R.; Stukenberg, P. T.; O'Donnell, M. *J. Biol. Chem.*, **1991**, *266*, 21681-6.

[138] Kelman, Z.; Yuzhakov, A.; Andjelkovic, J.; O'Donnell, M. *Embo J.*, **1998**, *17*, 2436-49.

[139] Yuzhakov, A.; Kelman, Z.; O'Donnell, M. *Cell*, **1999**, *96*, 153-63.

[140] O'Donnell, M.; Studwell, P. S. *J. Biol. Chem.*, **1990**, *265*, 1179-87.

[141] Naktinis, V.; Onrust, R.; Fang, L.; O'Donnell, M. *J. Biol. Chem.*, **1995**, *270*, 13358-65.

[142] Guenther, B.; Onrust, R.; Sali, A.; O'Donnell, M.; Kuriyan, J. *Cell*, **1997**, *91*, 335-45.

[143] Lee, S. H.; Walker, J. R. *Proc. Natl. Acad. Sci. USA*, **1987**, *84*, 2713-7.

[144] Tsuchihashi, Z.; Kornberg, A. *J. Biol. Chem.*, **1989**, *264*, 17790-5.

[145] Hingorani, M. M.; O'Donnell, M. *J. Biol. Chem.*, **1998**, *273*, 24550-63.

[146] Hingorani, M. M.; Bloom, L. B.; Goodman, M. F.; O'Donnell, M. *Embo J.*, **1999**, *18*, 5131-44.

[147] Ason, B.; Bertram, J. G.; Hingorani, M. M.; Beechem, J. M.; O'Donnell, M.; Goodman, M. F.; Bloom, L. B. *J. Biol. Chem.*, **2000**, *275*, 3006-15.

[148] Mossi, R.; Hubscher, U. *Eur. J. Biochem.*, **1998**, *254*, 209-16.

[149] Cai, J.; Yao, N.; Gibbs, E.; Finkelstein, J.; Phillips, B.; O'Donnell, M.; Hurwitz, J. *Proc. Natl. Acad. Sci. USA*, **1998**, *95*, 11607-12.

[150] Noskov, V.; Maki, S.; Kawasaki, Y.; Leem, S. H.; Ono, B.; Araki, H.; Pavlov, Y.; Sugino, A. *Nucleic Acids Res.*, **1994**, *22*, 1527-35.

[151] Li, X.; Burgers, P. M. *Proc. Natl. Acad. Sci. USA*, **1994**, *91*, 868-72.

[152] Jarvis, T. C.; Paul, L. S.; von Hippel, P. H. *J. Biol. Chem.*, **1989**, *264*, 12709-16.

[153] Berdis, A. J.; Benkovic, S. J. *Biochemistry*, **1996**, *35*, 9253-65.

[154] Kaboord, B. F.; Benkovic, S. J. *Curr. Biol.*, **1995**, *5*, 149-57.

[155] Young, M. C.; Weitzel, S. E.; von Hippel, P. H. *J. Mol. Biol.*, **1996**, *264*, 440-52.

[156] Sexton, D. J.; Carver, T. E.; Berdis, A. J.; Benkovic, S. J. *J. Biol. Chem.*, **1996**, *271*, 28045-51.

[157] Latham, G. J.; Dong, F.; Pietroni, P.; Dozono, J. M.; Bacheller, D. J.; von Hippel, P. H. *Proc. Natl. Acad. Sci. USA*, **1999**, *96*, 12448-53.

[158] Alley, S. C.; Abel-Santos, E.; Benkovic, S. J. *Biochemistry*, **2000**, *39*, 3076-90.

[159] Sexton, D. J.; Kaboord, B. F.; Berdis, A. J.; Carver, T. E.; Benkovic, S. J. *Biochemistry*, **1998**, *37*, 7749-56.

[160] Lee, S. H.; Hurwitz, J. *Proc. Natl. Acad. Sci. USA*, **1990**, *87*, 5672-6.

[161] Burgers, P. M. *J. Biol. Chem.*, **1991**, *266*, 22698-706.

[162] Cai, J.; Gibbs, E.; Uhlmann, F.; Phillips, B.; Yao, N.; O'Donnell, M.; Hurwitz, J. *J. Biol. Chem.*, **1997**, *272*, 18974-81.

[163] Uhlmann, F.; Cai, J.; Gibbs, E.; O'Donnell, M.; Hurwitz, J. *J. Biol. Chem.*, **1997**, *272*, 10058-64.

[164] Fotedar, R.; Mossi, R.; Fitzgerald, P.; Rousselle, T.; Maga, G.; Brickner, H.; Messier, H.; Kasibhatla, S.; Hubscher, U.; Fotedar, A. *Embo J.*, **1996**, *15*, 4423-33.

[165] Ellison, V.; Stillman, B. *J. Biol. Chem.*, **1998**, *273*, 5979-87.

[166] Maga, G.; Mossi, R.; Fischer, R.; Berchtold, M. W.; Hubscher, U. *Biochemistry*, **1997**, *36*, 5300-10.

[167] Song, Q.; Lu, H.; Zhang, N.; Luckow, B.; Shah, G.; Poirier, G.; Lavin, M. *Biochem Biophys Res Commun*, **1997**, *233*, 343-8.

[168] Rheaume, E.; Cohen, L. Y.; Uhlmann, F.; Lazure, C.; Alam, A.; Hurwitz, J.; Sekaly, R. P.; Denis, F. *Embo J.*, **1997**, *16*, 6346-54.

[169] Tougu, K.; Peng, H.; Marians, K. J. *J. Biol. Chem.*, **1994**, *269*, 4675-82.

[170] Naktinis, V.; Turner, J.; O'Donnell, M. *Cell*, **1996**, *84*, 137-45.

[171] Murakami, Y.; Eki, T.; Hurwitz, J. *Proc. Natl. Acad. Sci. USA*, **1992**, *89*, 952-6.

[172] Fukuda, K.; Morioka, H.; Imajou, S.; Ikeda, S.; Ohtsuka, E.; Tsurimoto, T. *J. Biol. Chem.*, **1995**, *270*, 22527-34.

[173] Warbrick, E.; Lane, D. P.; Glover, D. M.; Cox, L. S. *Oncogene*, **1997**, *14*, 2313-21.

[174] Gary, R.; Ludwig, D. L.; Cornelius, H. L.; MacInnes, M. A.; Park, M. S. *J. Biol. Chem.*, **1997**, *272*, 24522-9.

[175] Levin, D. S.; Bai, W.; Yao, N.; O'Donnell, M.; Tomkinson, A. E. *Proc. Natl. Acad. Sci. USA*, **1997**, *94*, 12863-8.

[176] Chuang, L. S.; Ian, H. I.; Koh, T. W.; Ng, H. H.; Xu, G.; Li, B. F. *Science,* **1997**, *277*, 1996-2000.

[177] Warbrick, E. *Bioessays,* **1998**, *20*, 195-9.

[178] Kelman, Z.; Hurwitz, J. *Trends Biochem. Sci.,* **1998**, *23*, 236-8.

[179] Lieber, M. R. *Bioessays,* **1997**, *19*, 233-40.

[180] Flores-Rozas, H.; Kelman, Z.; Dean, F. B.; Pan, Z. Q.; Harper, J. W.; Elledge, S. J.; O'Donnell, M.; Hurwitz, J. *Proc. Natl. Acad. Sci. USA,* **1994**, *91*, 8655-9.

[181] Johnson, R. E.; Prakash, S.; Prakash, L. *Science,* **1999**, *283*, 1001-4.

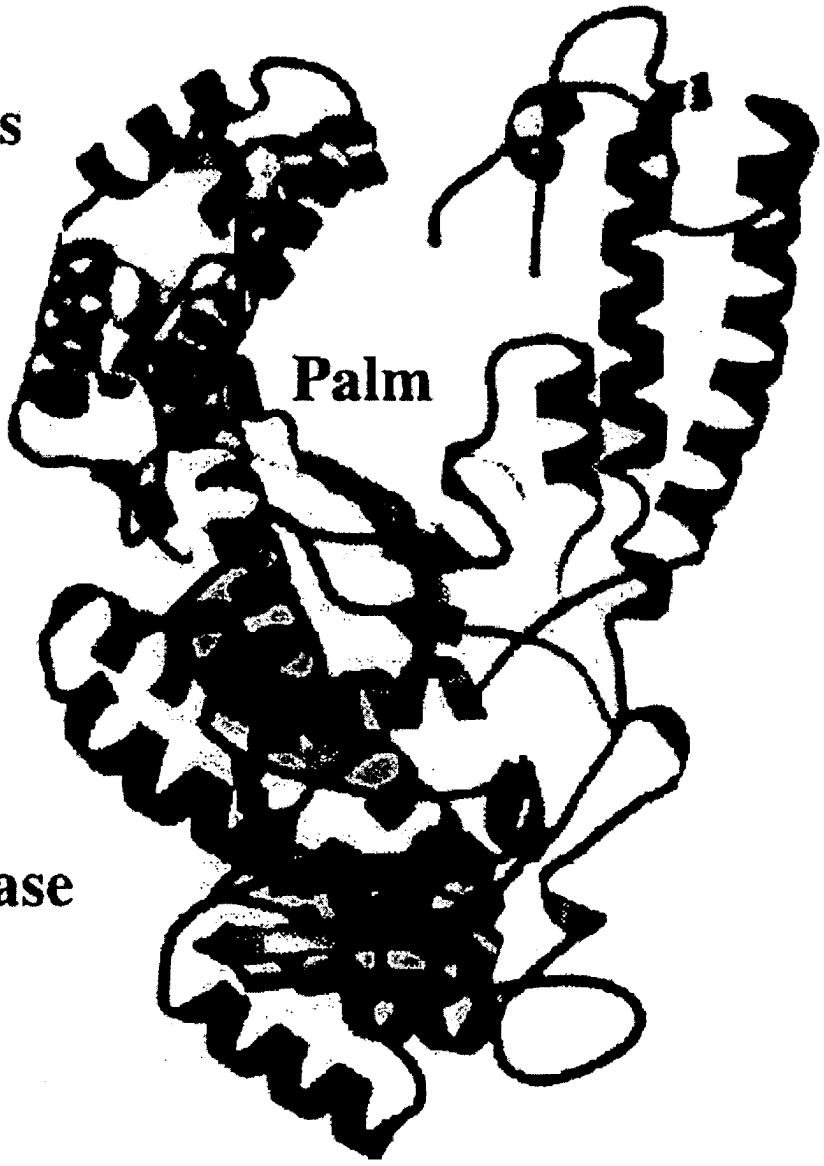
[182] Washington, M. T.; Johnson, R. E.; Prakash, S.; Prakash, L. *J. Biol. Chem.,* **1999**, *274*, 36835-8.

[183] Washington, M. T.; Johnson, R. E.; Prakash, S.; Prakash, L. *Proc. Natl. Acad. Sci. USA,* **2000**, *97*, 3094-9.

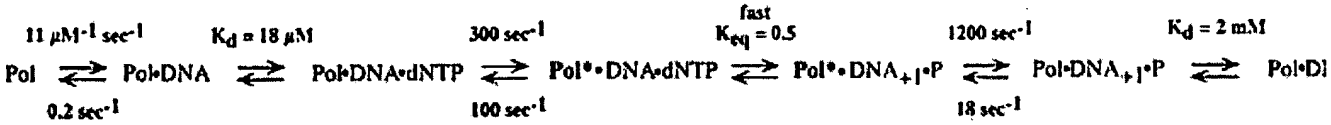
[184] Nelson, J. R.; Lawrence, C. W.; Hinkle, D. C. *Science,* **1996**, *272*, 1646-9.

Function	Bacteria (*E. coli*)	Bacteriophage (T4)	Eukarya (humans / *S. cerevisiae*)	Archaea
Helicase	DnaB (homohexamer)	gp41 (homohexamer)	MCM (6 subunits)	MCM homolog (1 to 4 subunits)
Primase	DnaG	gp61	Pol α/primase (4 subunits)	Primase homolog
Single-strand DNA binding protein	SSB	gp32	RPA (3 subunits)	RPA homolog
Polymerase (3'-5' exonuclease)	core polymerase (3 subunits: α,ε,θ)	gp43	Pol δ (3-4 subunits) Polε? (_ 5 subunits)	Family B polymerases DP2-like polymerases?
Sliding clamp	β (dimer)	gp45 (trimer)	PCNA (trimer)	PCNA homolog (trimer)
Clamp loader	γ complex (γδδ' χψ)	gp44/62 (4 gp44 + 1 gp62)	RFC (5 subunits)	RFC homolog (two subunits)

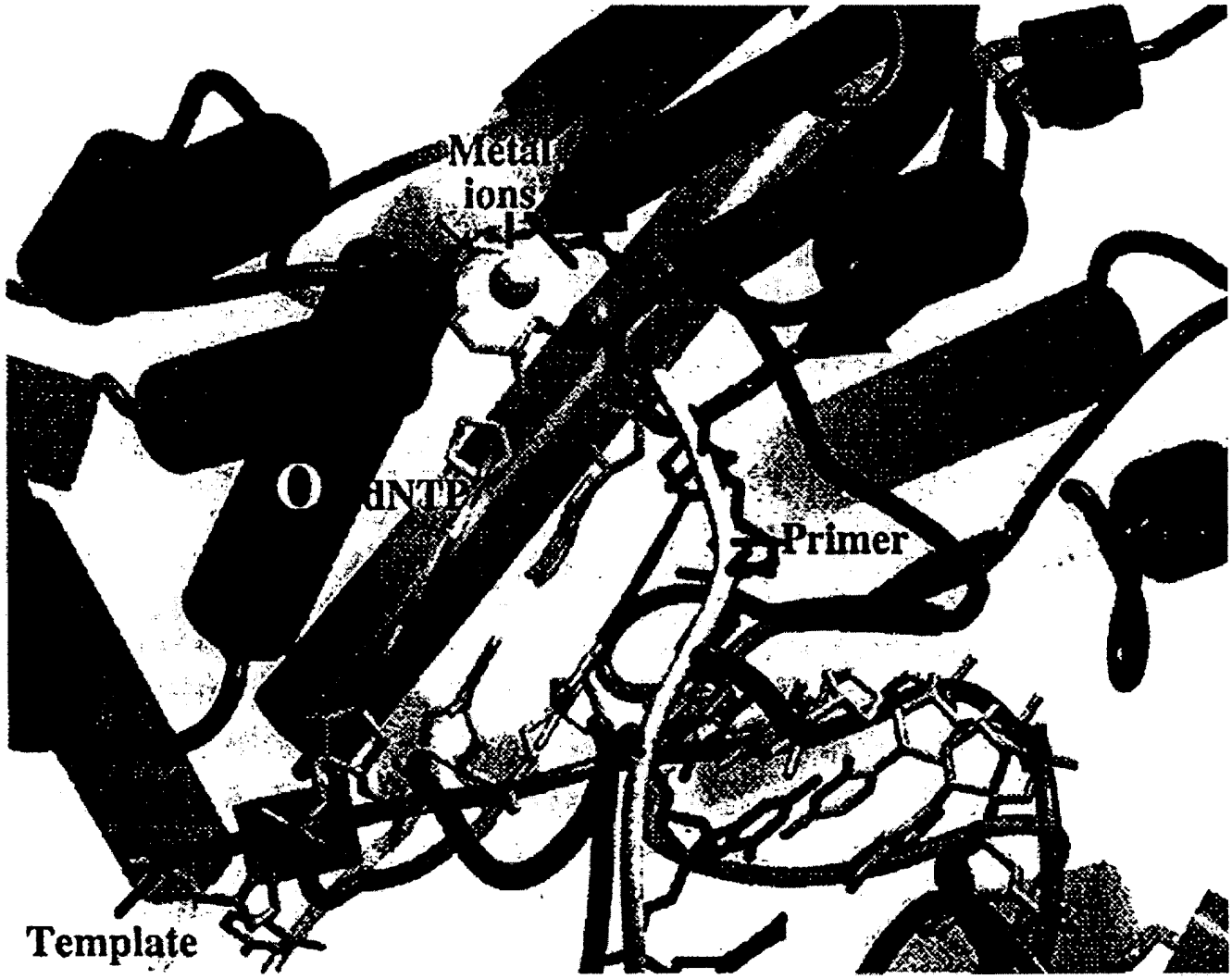
E. coli Pol I Klenow Fragment



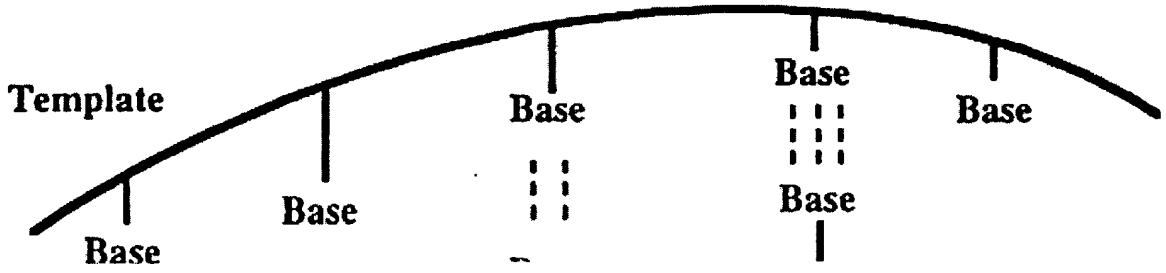
T7 DNA polymerase mechanism



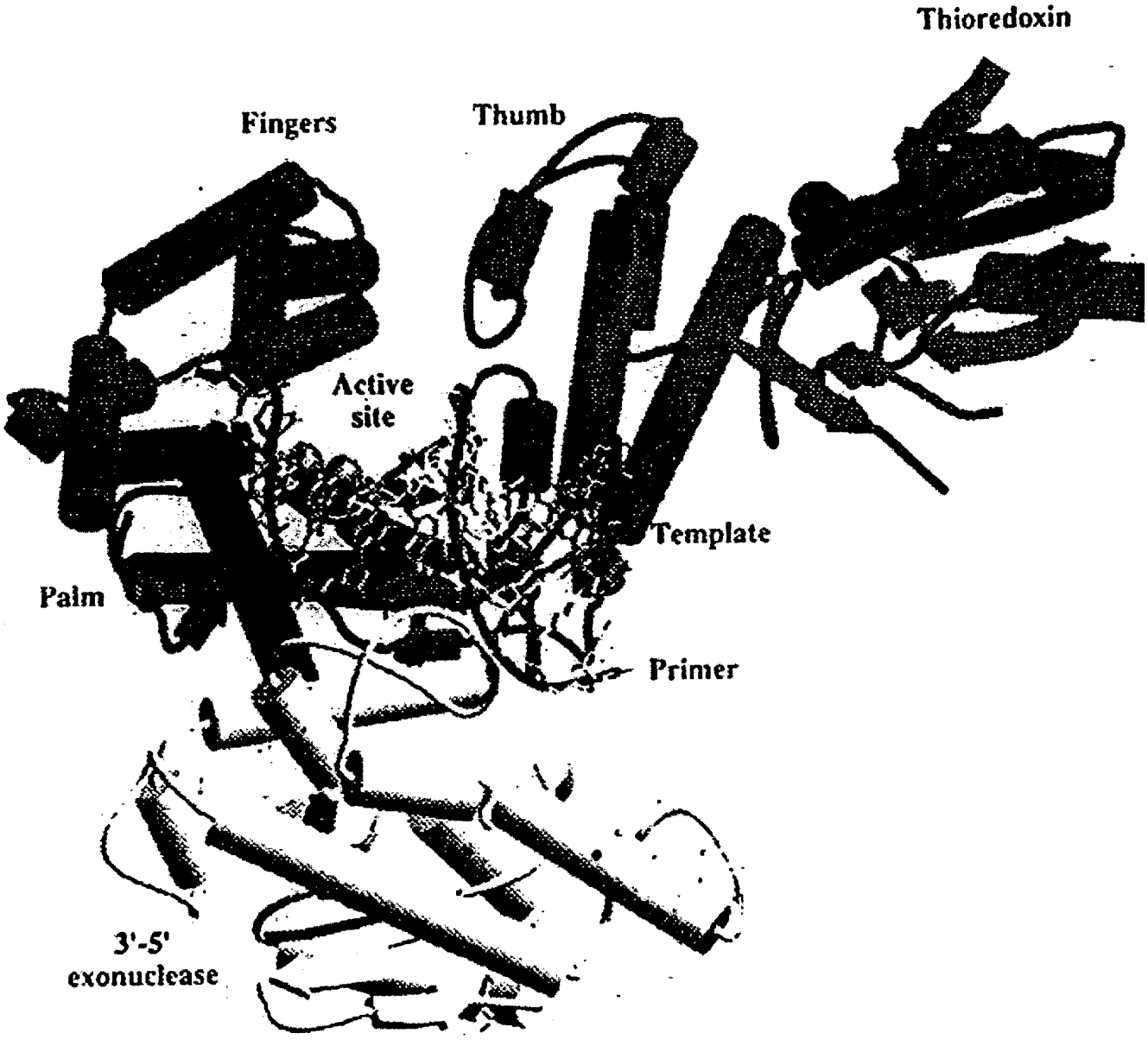
A. T7 DNA polymerase active site

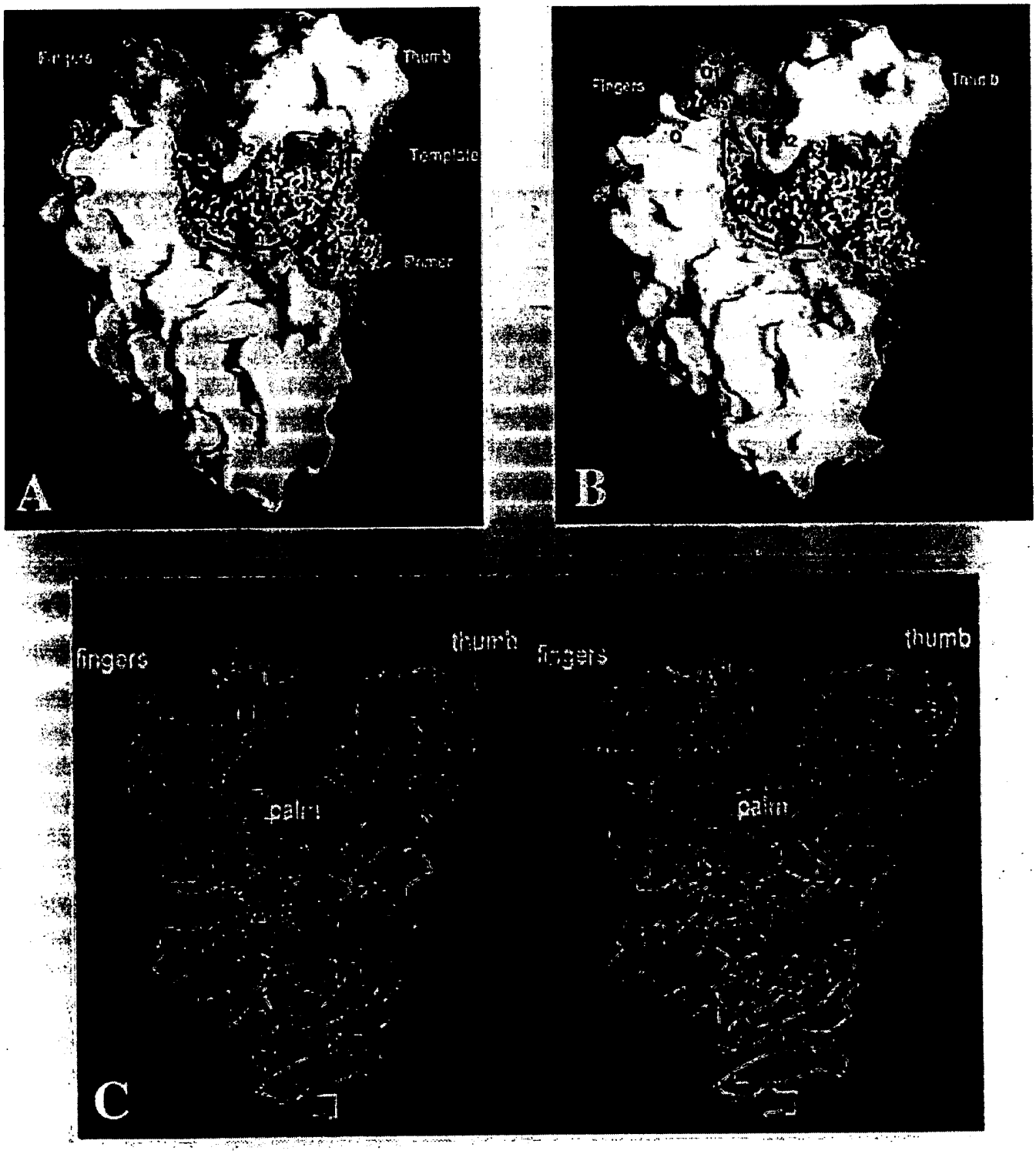


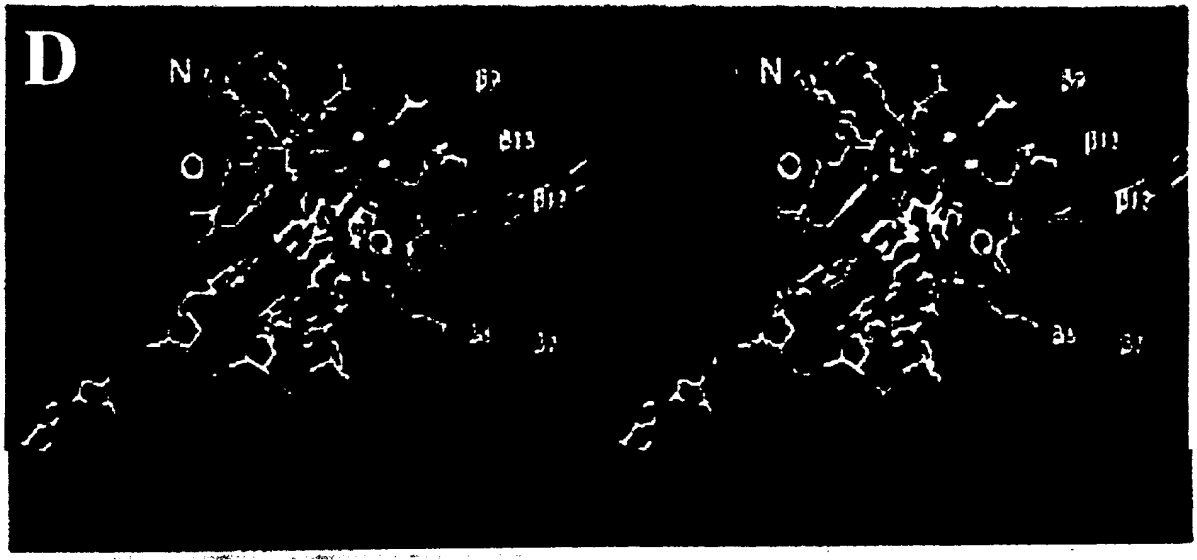
B. Two metal ion mechanism for nucleotidyl transfer

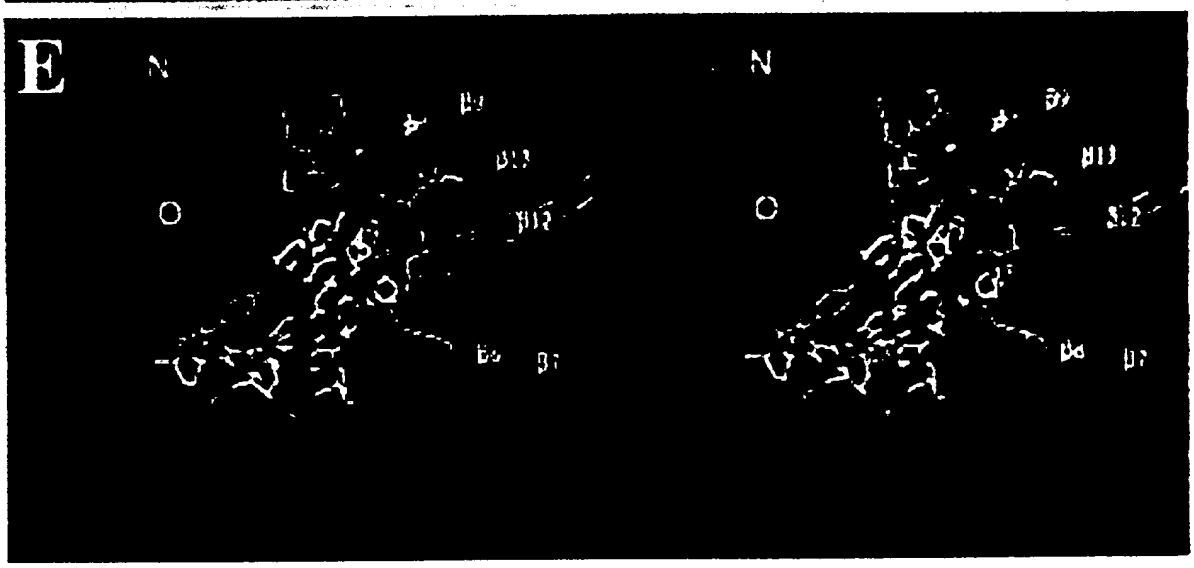


T7 DNA polymerase with thioredoxin



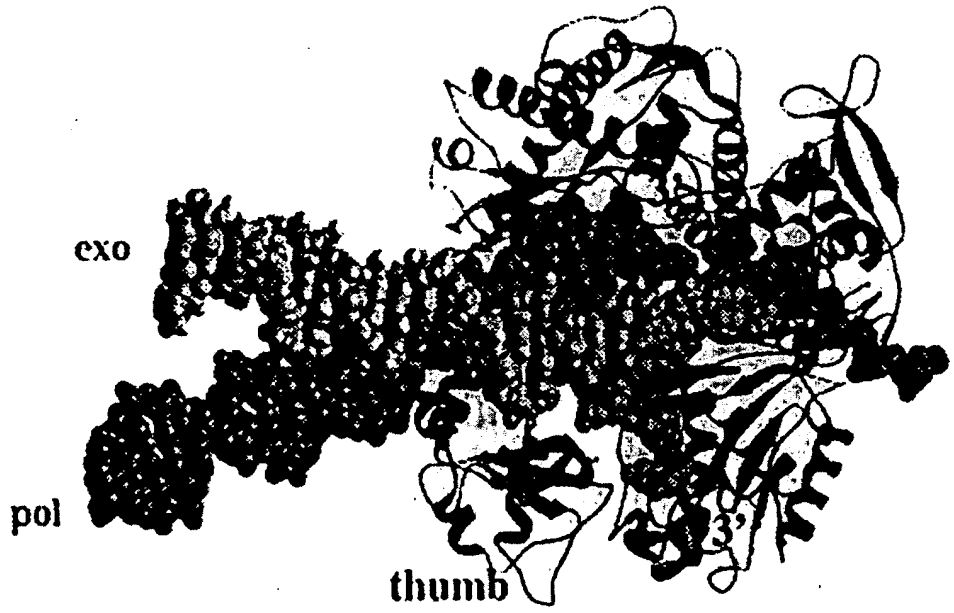

Fingers
Thumb
Template
Primer
A
Fingers
Thumb
B
fingers
thumb
palm
fingers
thumb
palm
C


D
N
O
N
O

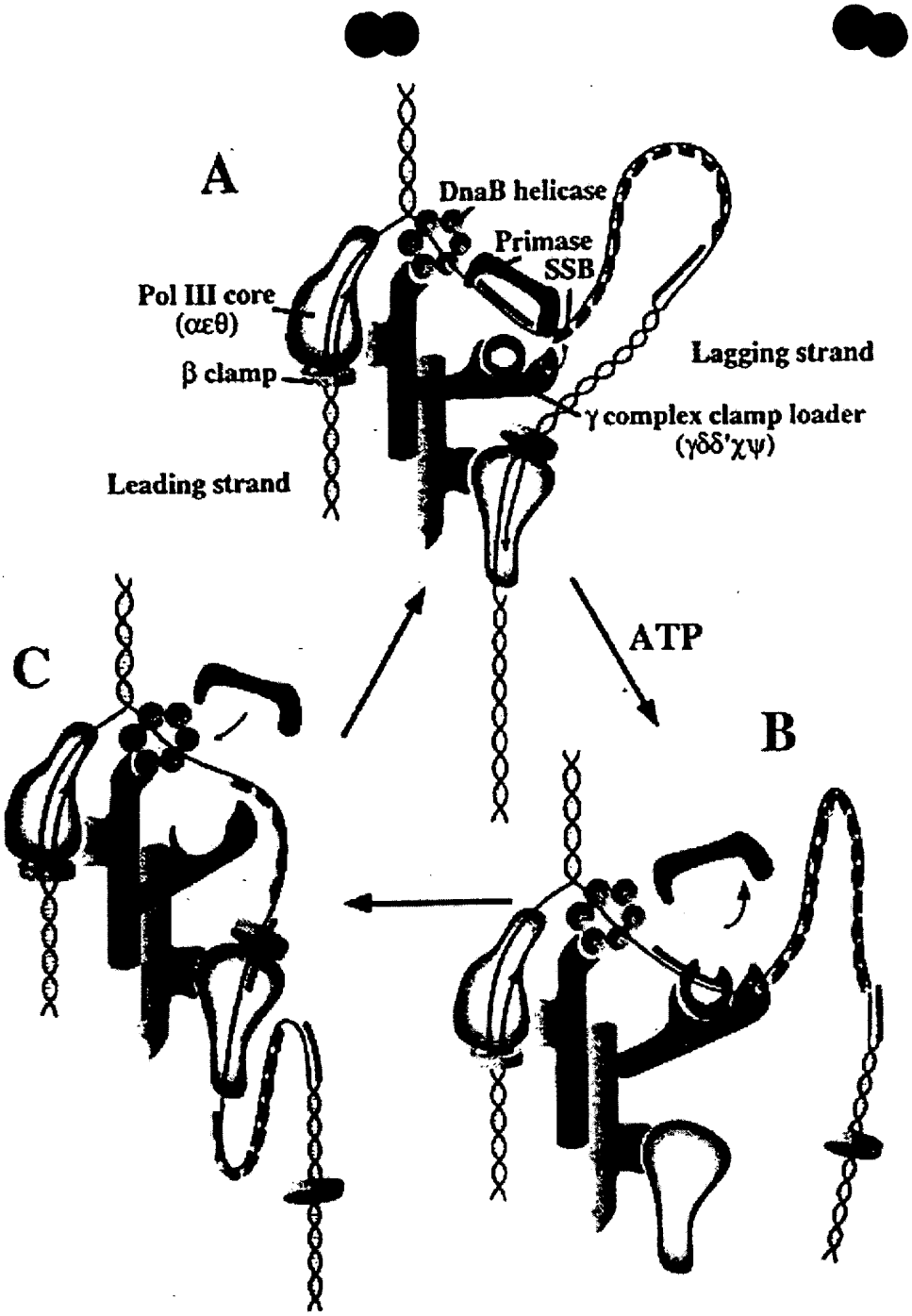

E
N
O
N
O

A. RB69 polymerase with DNA at the editing site

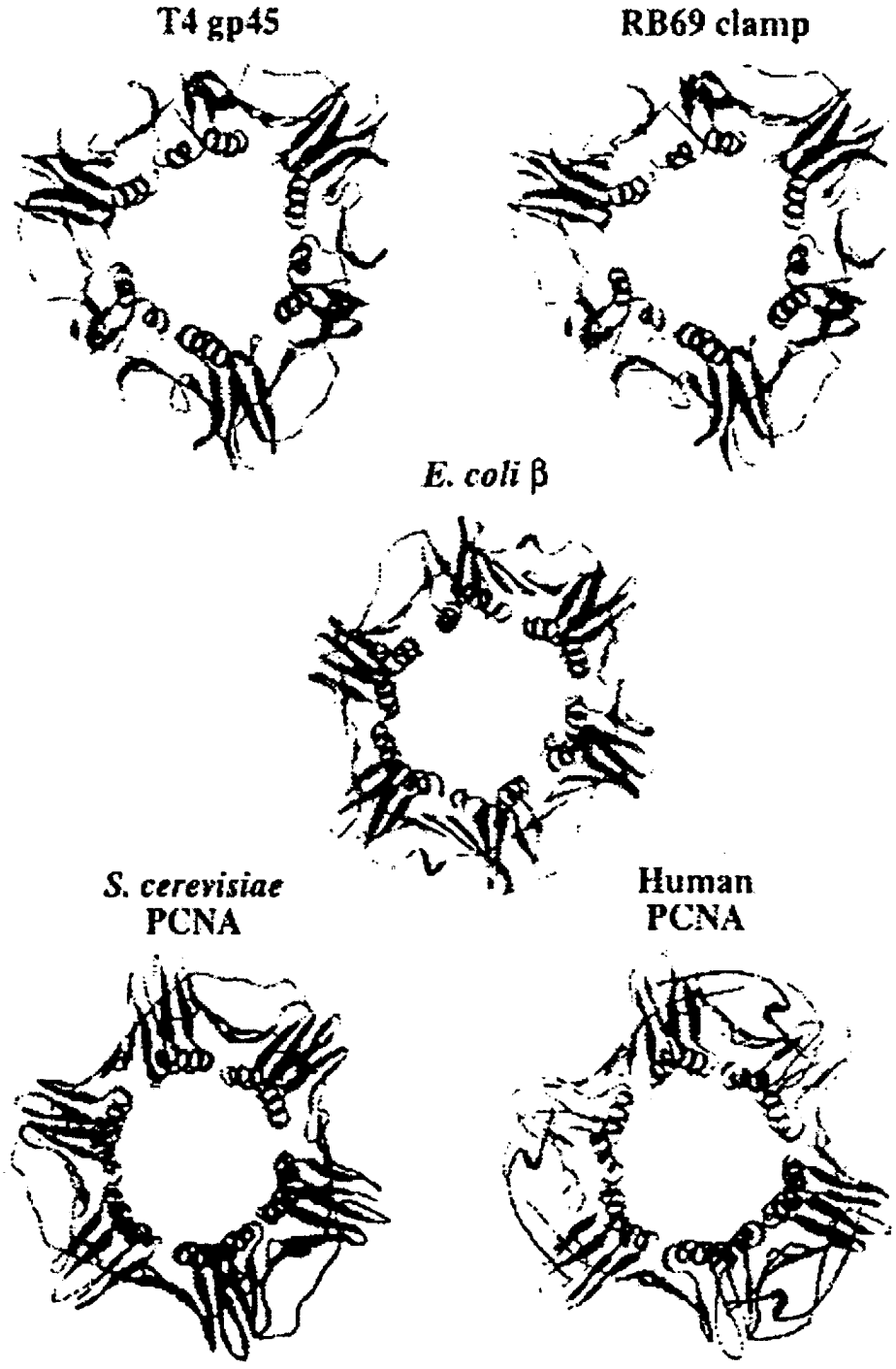
B. Polymerase *versus* editing mode



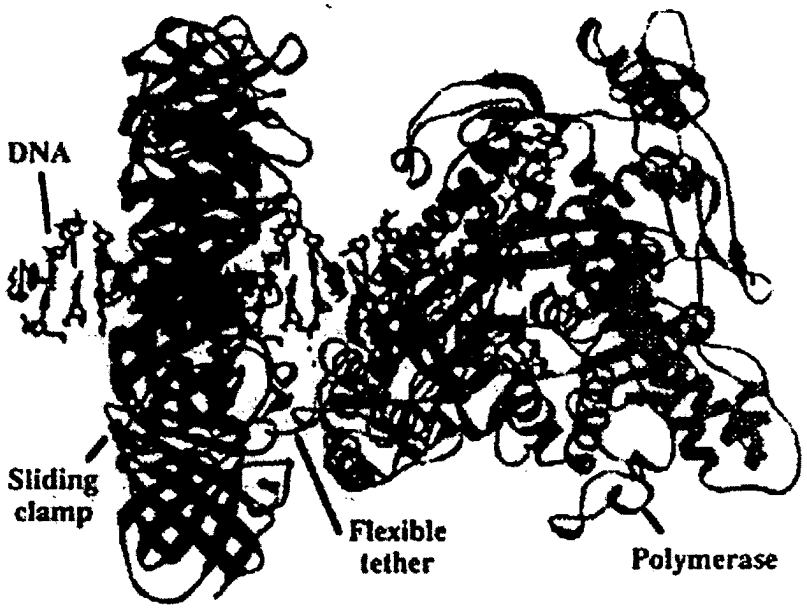
http://www.bentham.org/coc-sample/o-donnell/fig6.jpg


A
DnaB helicase
Primase
SSB
Pol III core
(αεθ)
β clamp
Lagging strand
γ complex clamp loader
(γδδ'χψ)
Leading strand
ATP
B
C

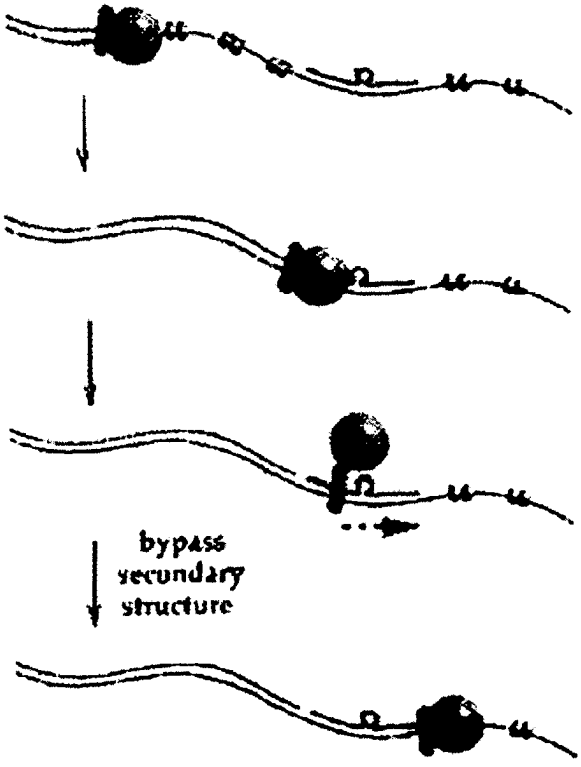
A. Sliding clamps



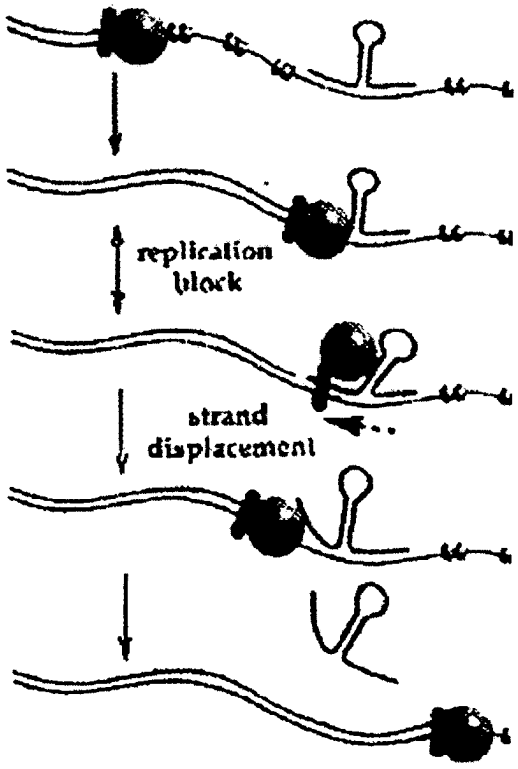
B. RB69 polymerase with clamp on DNA

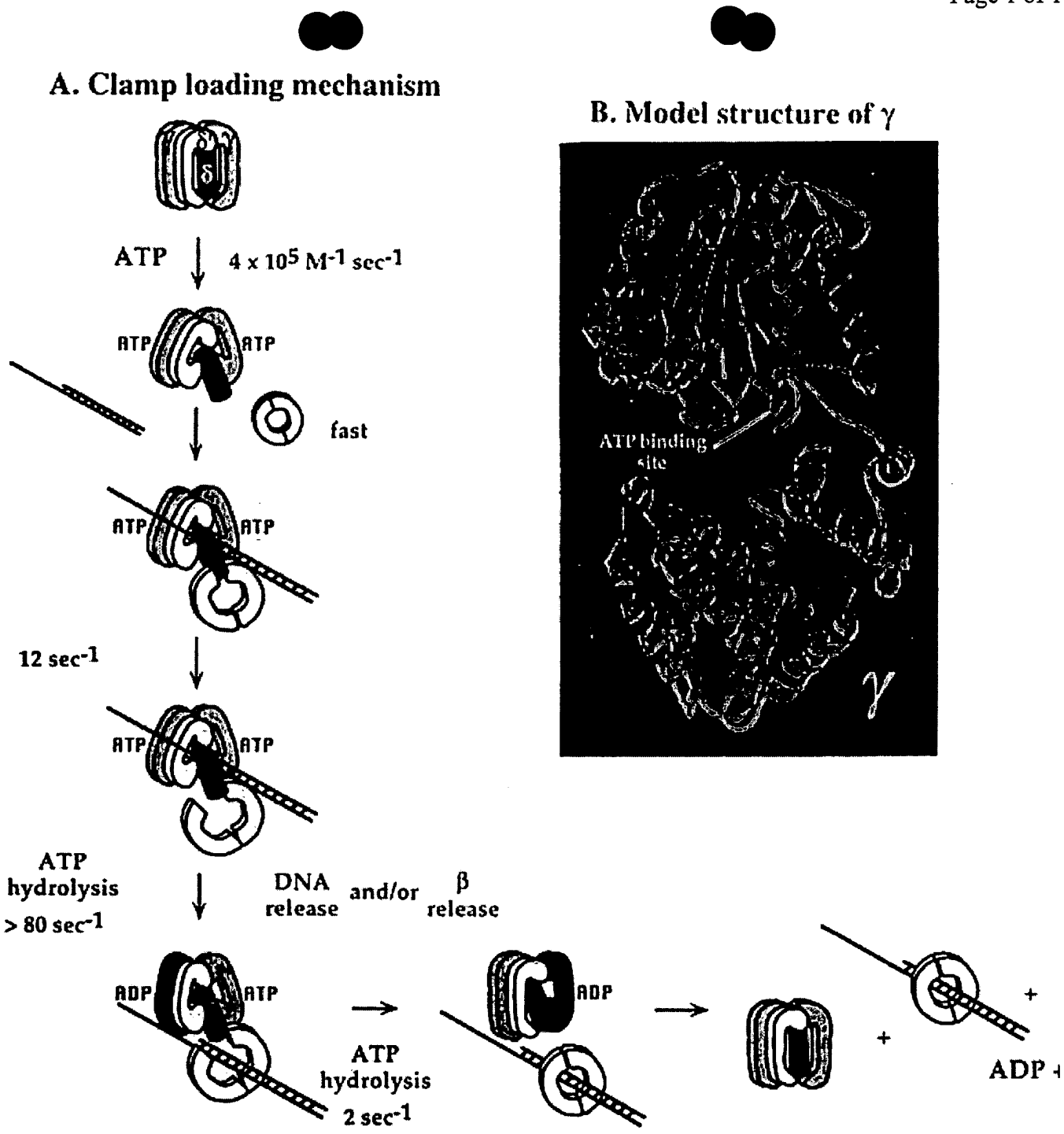


C. Strand traversal



D. Strand displacement




A. Clamp loading mechanism
B. Model structure of γ
δ
ATP
4 x 10^5 M^-1 sec^-1
ATP
ATP
fast
ATP
ATP
12 sec^-1
ATP
ATP
ATP hydrolysis
> 80 sec^-1
DNA release
and/or
β release
ADP
ATP
ATP hydrolysis
2 sec^-1
ADP
ADP +
ATP binding site
γ